RTO-MP-058
AC/323(HFM-058)TP/30

RTO-MP-058

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 58

What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?

(les Caractéristiques essentielles des systèmes VR pour atteindre les objectifs militaires en matière de performances humaines)

Papers presented at the RTO Human Factors and Medicine Panel (HFM) Workshop held in The Hague, The Netherlands, 13-15 April 2000.

Published March 2001

Distribution and Availability on Back Cover

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY) 01-03-2001	**2. REPORT TYPE** RTO Meeting Proceedings	**3. DATES COVERED (FROM - TO)** xx-xx-2001 to xx-xx-2001
4. TITLE AND SUBTITLE What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals? Unclassified		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME AND ADDRESS Research and Technology Organization North Atlantic Treaty Organization BP 25, 7 rue Ancelle F-92201 Neuilly-sur-Seine Cedex , France F-92201		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS ,		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT A PUBLIC RELEASE		

,
13. SUPPLEMENTARY NOTES
14. ABSTRACT This workshop aimed to identify the functional requirements of potential military applications of Virtual Reality (VR) technology, to report the state-of-the-art and projected capabilities of VR technologies, and to propose future research requirements and directions for military applications. During the workshop discussions, forty participants from military organisations, academia and industry put forward their opinions on the significant bottlenecks and opportunities in the development of military VR applications. Presentations discussed visual, haptic, auditory and motion feedback, navigation interfaces, and scenario generation, modelling software and rendering hardware. VR research transition opportunities include the domains of training, planning & mission rehearsal, simulation supported operation, remotely operated systems and product design. Critical bottlenecks are a lack of natural interfaces, a lack of technology standardisation and a lack of behavioural models and team interaction tools in VR. In general, better co-ordination between military organisations, industry and academia is necessary in order to identify gaps in current knowledge and to co-ordinate research. Suggestions for closing gaps are included.
15. SUBJECT TERMS Military training; human factors engineering; user needs; interfaces; research projects; motion sickness; similators; performance; humans; man machine systems; perception; virtual reality; virtual environments

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
			Public Release	172	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified			**19b. TELEPHONE NUMBER** **International Area Code** **Area Code Telephone Number** 703 767-9007 **DSN** 427-9007

This page has been deliberately left blank

Page intentionnellement blanche

RTO-MP-058
AC/323(HFM-058)TP/30

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 58

What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?

(les Caractéristiques essentielles des systèmes VR pour atteindre les objectifs militaires en matière de performances humaines)

Papers presented at the RTO Human Factors and Medicine Panel (HFM) Workshop held in The Hague, The Netherlands, 13-15 April 2000.

The Research and Technology Organization (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote cooperative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective coordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also coordinates RTO's cooperation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of initial cooperation.

The total spectrum of R&T activities is covered by the following 7 bodies:

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS Studies, Analysis and Simulation Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier cooperation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced
directly from material supplied by RTO or the authors.

Published March 2001

ISBN 92-837-1057-6

Printed by St. Joseph Ottawa/Hull
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada J8X 1C6

What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?

(RTO MP-058 / HFM-058)

Executive Summary

PURPOSE

The purpose of the workshop was to:

- identify the functional requirements of potential military applications of Virtual Reality (VR) technology,
- report the state-of-the-art and projected capabilities of VR technologies, and
- propose future research requirements and directions for military applications.

SUMMARY

The workshop was organised into three daylong sessions. The first day focused on functional requirements for military VR applications in the domains of training, robotics, remote operations and command and control. On the second day, we examined available VR techniques now and in the near future. Presentations discussed visual, haptic, auditory and motion feedback, navigation interfaces, and scenario generation, modelling software and rendering hardware. The third day addressed missing VR capability and future research and concluded with a panel discussion.

During the workshop discussions forty participants from military organisations, academia and industry put forward their opinions on the biggest bottlenecks and opportunities in the development of military VR applications.

MAIN CONCLUSIONS

Virtual Reality technology is of great interest to the military. Its most important application domain is training. VR for training can reduce cost and risk of casualties and improve flexibility and performance monitoring. Furthermore, great opportunities are identified in the domains of planning and mission rehearsal, simulation supported operation, remotely operated systems and product design.

At the same time a number of factors seem to frustrate successful applications in this field. One of the significant bottlenecks is that VR developments are usually not user driven. Application developers and designers do not pay enough attention to human factors requirements. Consequently, applications may fail because of a lack of natural interfaces and motion sickness. So far, user interfaces have been poorly attuned to natural human skills (crude input devices and inconsistent visual, auditory and proprioceptive feedback) and to the tasks to be performed in VR. A second bottleneck is the lack of standardisation causing problems with integrating VR systems and VR software tools. A third is the lack of behavioural models of people and objects in VR scenarios and facilities for team interactions (poor visual human representations and communication tools).

MAJOR RECOMMENDATIONS

In general, better co-ordination between military organisations, industry and academia is essential in order to identify gaps in current knowledge and co-ordinate research. To this purpose the military should develop a vision on the use of VR technology and specify their needs more clearly. Industry should work on standardisation and should substantially implement human factors into their development process. Academia and research institutes should co-ordinate and accelerate their long-term research efforts to focus on natural interfaces (innovative metaphors) and on how to model (intelligent) human and object behaviour. In the short term academia should focus on human factors metrics and metrics for team performance (cognition, communication), and a standard evaluation methodology.

A specific suggestion made during the workshop that could contribute to solving the bottlenecks is to establish a RTO Task Group to (1) identify applications with a high return of investment, user requirements and technologies for investment by the military and (2) foster development of natural VR interfaces and behaviourally realistic intelligent agents and models (identify new funding sources).

The enthusiasm of the workshop attendees and the evident willingness to share ideas and to discuss their findings provide a promising base for a co-operation between military agencies, industry and academia. Research on the usability of VR technology will enable militaries to be smart buyers. It will ensure that Virtual Reality hardware and software is capable of meeting the perceptual, fidelity, transfer of training, and health and safety requirements of applications.

les Caractéristiques essentielles des systèmes VR pour atteindre les objectifs militaires en matière de performances humaines

(RTO MP-058 / HFM-058)

Synthèse

OBJET

L'atelier avait pour objet :

- d'identifier les besoins fonctionnels découlant des applications militaires possibles des technologies de réalité virtuelle (VR),
- de rendre compte de l'état actuel des connaissances et des capacités anticipées dans ce domaine, et
- de proposer de futurs sujets de recherche et des orientations vers des applications militaires.

RÉSUMÉ

L'atelier a été organisé en trois sessions d'une journée : La première journée a été consacrée aux besoins fonctionnels découlant des applications militaires des technologies VR dans les domaines de l'entraînement, la robotique, les opérations à distance et le contrôle. Le deuxième jour, nous avons examiné les techniques VR actuelles et émergentes. Des présentations ont été données sur le bouclage de l'information dans les domaines visuels, haptiques, auditifs, et cybernétiques, les interfaces de navigation, la génération de scénarios, les logiciels de modélisation et le matériel de rendu d'image. La troisième journée a été centrée sur les capacités faisant défaut dans le domaine de la VR, ainsi que les travaux de recherche futurs, et s'est terminée par une discussion entre les membres de la commission.

Au cours des discussions qui ont eu lieu pendant les trois jours de l'atelier, une quarantaine de participants venus d'organisations militaires, d'universités et de l'industrie ont exprimé leurs opinions sur les impasses les plus importantes, ainsi que sur les opportunités offertes de développer de nouvelles applications VR militaires.

CONCLUSIONS PRINCIPALES

Les technologies de réalité virtuelle sont d'un grand intérêt pour les militaires. Le domaine d'application le plus important est celui de l'entraînement. L'emploi de techniques VR pour l'entraînement permettrait de réduire son coût, ainsi que le risque d'accidents corporels, et pourrait apporter des améliorations au niveau de la flexibilité et du contrôle des performances. En outre, de grandes possibilités ont déjà été identifiées dans les domaines de la planification et la préparation des missions, de la conduite des opérations à l'aide de la simulation, de la télécommande des systèmes et de la conception des produits.

En même temps, un certain nombre de facteurs sembleraient entraver la réussite des applications dans ce domaine. Le fait que les développements en matière de VR soient rarement orientés par les utilisateurs représente l'une des principales gênes. Les développeurs d'applications et les concepteurs ne tiennent pas suffisamment compte des besoins du point de vue des facteurs humains. Par conséquent, les applications risquent d'échouer du fait du mal des transports et du manque d'interfaces naturelles. Jusqu'à présent, les interfaces utilisateurs ont été mal adaptées aux capacités humaines naturelles (des unités d'entrée rustiques et des boucles d'information visuelles, auditives et proprioceptives incompatibles) ainsi qu'aux tâches à accomplir en VR. Le manque de normalisation, qui crée des problèmes d'intégration des systèmes et des outils VR représente une deuxième gêne importante. Enfin, le manque de modèles du comportement humain et d'objets dans les scénarios VR, ainsi que le manque de possibilités d'interactions interéquipes (représentations visuelles du corps humain et outils de communication de mauvaise qualité) est la troisième gêne identifiée.

RECOMMANDATIONS PRINCIPALES

De façon générale, il est indispensable d'assurer une meilleure coordination entre les organisations militaires, l'industrie et les universités, afin d'identifier les éventuelles lacunes dans les connaissances et de coordonner les travaux de recherche. Avec cet objectif en vue, les militaires devraient élaborer une philosophie de mise en oeuvre des technologies VR et exprimer leurs besoins plus clairement. L'industrie devrait travailler sur la normalisation et faire une large place aux facteurs humains dans leurs processus de développement. Les universités et les instituts de recherche devrait coordonner et intensifier leurs efforts de recherche à long terme afin de se concentrer sur les interfaces naturelles (métaphores novatrices) et sur la modélisation (intelligente) du comportement des objets et des êtres humains. A court terme, les universitaires devraient privilégier la métrologie des facteurs humains et la métrologie du travail en équipe (l'approche cognitive, la communication), ainsi que l'élaboration d'une nouvelle méthodologie normalisée d'évaluation.

L'une des propositions faites au cours de l'atelier, qui pourrait contribuer à l'élimination de impasses, consisterait à créer un groupe de travail RTO pour (1) identifier des applications ayant un bon rendement, les besoins des utilisateurs et les technologies méritant des efforts d'investissement de la part des militaires, et (2) encourager le développement d'interfaces VR naturelles, ainsi que des agents et des modèles intelligents ayant des comportements réalistes (identification de nouveaux bailleurs de fonds).

L'enthousiasme manifesté par les participants durant l'atelier, ainsi que leur volonté évidente de partager leurs idées et de discuter de leurs conclusions a constitué une base prometteuse pour une coopération future entre les agences militaires, l'industrie et les universités. Des recherches doivent être entreprises sur la facilité d'utilisation de ces technologies afin de permettre aux militaires de les acheter en connaissance de cause. Ils pourraient ainsi s'assurer que le matériel et les logiciels de réalité virtuelle seraient compatibles avec les exigences de perception, de fidélité, de transfert d'entraînement et d'hygiène et sécurité demandées pour les applications.

Click inside the blue boxes or on the titles to view the corresponding section

(except for items marked in red, which were not available at the time of production)

Contents

† Paper not available at time of printing.

† Paper not available at time of printing.

Human Factors and Medicine Panel

Chairman:

Dr M.C. WALKER
Director, Centre for Human Sciences
DERA
F138 Building - Room 204
Farnborough, Hants GU14 0LX
United Kingdom
Tel: 44 1252 393 764
Fax: 44 1252 393 982
Email: mcwalker@dera.gov.uk

Vice-Chairman:

Col. W. C. M. TIELEMANS, MD
RNLAF/SGO
P.O. Box 20703
Binckhorstlaan, 135
2500 ES The Hague
The Netherlands
Tel.: 31 70 339 6403
Fax: 31 70 339 7439
Email: willem.cobina.tielemans@planet.nl

PROGRAMME COMMITTEE

Workshop Chairman

Dr. Peter WERKHOVEN
TNO Human Factors Research Institute
Dept. of Work Environment
Kampweg 5
3769 ZG Soesterberg, The Netherlands
Tel.: +31 3463 56283 Fax: +31 3463 53977
Email: werkhoven@tm.tno.nl

Members

Thomas ALEXANDER
FGAN/FKIE
Neuenahrer Str. 20
53343 Wachtberg-Werthhoven, Germany
Tel.: +49 (0)228 9435 480
Fax: +49 (0)228 9435 508
Email: alex@fgan.de

Dr. Robert BREAUX
NAWC-TSD
12350 Research Parkway
Orlando, Florida 32826
United States of America
Tel.: +1 407 380 8168
Fax: +1 407 380 4007
Email: BreauxRB@navair.navy.mil

Dr. Stephen L. GOLDBERG
US Army Research Institute
12350 Research Parkway
Orlando, Florida 32826-3276
United States of America
Tel.: +1 407 384 3980
Fax: +1 407 384 3999
Email: Stephen_Goldberg@stricom.army.mil

Pascal HUE
DGA-ETAS
BP36
49460 Montreuil Juigne, France
Tel.: +33 241936644
Fax: +33 241936704
Email: huepascal@compuserve.com

Dr. Martin G. KAYE
DERA Centre for Human Sciences
Room 2012, A5 Building
Farnborough, Hampshire, GU14 0LX
United Kingdom
Tel.: +44 1252 393610
Fax: +44 1252 394700
Email: mgkaye@dera.gov.uk

Trond MYHRER
Norwegian Defence Research Establishment
P.O.Box 25
N-2007 Kjeller
Norway
Tel.: +47 63 80 78 52
Fax: +47 63 80 78 11
Email: trond.myhrer@ffi.no

continued overleaf

Antonio GRAMAGE
ISDEFE
Edison, 4
28006 Madrid
Spain
Tel.: +34 1 4115011
Fax: +34 1 4114703
Email: agramaje@isdefe.es

Elizabeth HENDERSON
Department of Informatics and Simulation
Royal Military College of Science
Shrivenham, Swindon SN6 8LA
United Kingdom
Tel.: +44 1793 785652
Fax: +44 1793 782753
Email: hendereg@rmcs.cranfield.ac.uk

MCS Jean-Paul PAPIN
7, rue Roger
92140 CLAMART
France
Tel.: +33 141087317
Fax: /
email: j.papin@wanadoo.fr

Lisbeth M. RASMUSSEN
Danish Defence Research Establishment
Svanemollens Kasserne, Ryvangs Allé 1
DK-2100 Copenhagen OE
Denmark
Tel.: +45 39 15 18 05
Fax: +45 39 29 15 33
Email: lr@ddre.dk

PANEL EXECUTIVE

Dr C. WIENTJES
BP 25 - 7, Rue Ancelle
92201 Neuilly-sur-Seine, France
Tel: +33 1 55 61 22 60
Fax: +33 1 55 61 22 98
Email: wientjesc@rta.nato.int or pelatd@rta.nato.int

TECHNICAL EVALUATION REPORT

NATO WORKSHOP

"What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?"

The Hague, TNO-FEL, 13 – 15 April 2000

Peter Werkhoven
TNO Human Factors
Dept. of Information Processing
PO Box 23
3769 ZG Soesterberg, The Netherlands
werkhoven@tm.tno.nl

Robert Breaux
Naval Air Warfare Centre
Training Systems Division
12350 Research Parkway
Orlando, FL 32826-3275, USA
BreauxRB@navair.navy.mil

1 INTRODUCTION

1.1 Background (history HFM 021, previous activities)

NATO Research Study Group HFM-21 (called RSG-28 before the RTA reorganisation) was established to explore and evaluate human factors issues that effect the use of virtual reality technologies for military purposes. The findings of the group are to provide NATO countries with better understanding of the capabilities and limitations of this new and sometimes over-hyped technology. The study group has agreed upon the following definition of virtual reality to establish a common reference point.

> *Virtual Reality is the experience of being in a synthetic environment and the perceiving and interacting through sensors and effectors, actively and passively, with it and the objects in it, as if they were real. Virtual Reality technology allows the user to perceive and experience sensory contact and interact dynamically with such contact in any or all modalities.*

Virtual reality has great potential in areas such as training, mission rehearsal, concept development, weapon prototyping, and personnel selection. Many virtual reality technologies also have application in robotics and remote manipulation applications. There have been a number of successful research and prototype applications of virtual reality. They have mostly been in training. Use of virtual reality for ship-handling training has been successfully demonstrated in both the United States and Canada. Dismounted soldier simulation has seen considerable research and development activity. Virtual reality is a relatively new concept and many of the technologies involved in immersing individuals or teams in virtual environments are evolving and improving rapidly.

The key to the effectiveness of virtual reality for military purpose is the man-machine interface or human-computer interaction. Military personnel must be able to perform their tasks and missions using virtual reality sensory display devices and response devices. These devices must display an environment that provides the appropriate cues and responses needed to learn and perform military tasks. Human factors issues include: determining the perceptual capabilities and limitations of sensory display devices; designing terrain data bases and other displays to meet task performance needs; understanding the human and task performance compromises required by current technologies; evaluating transfer of training and knowledge from the virtual to the real world; and considering the causes and solutions to simulator sickness that can occur in virtual reality. The Research Study Group intends to provide information and recommendations on these issues to military researchers, requirements generators, and acquisition agencies. The intended benefit is better-informed decisions on application of virtual reality technologies to meet appropriate military needs.

Previous activities included:

- A one-day workshop titled “The development of a Generic Battery of Human Performance Metrics for Virtual Environments” (Chertsey, UK, October 14, 1996)
- A three-day workshop titled “The capability of Virtual Reality to meet military requirements” (Orlando FL, December 1997).

Reviews of these workshops together with a chapter on “Human Computer Interaction issues in VR” and a chapter on the “State of the art in VR research in NATO countries” will be published in the FINAL report of HFM-021.

The current workshop is the last activity organised by HFM-02:

- A three-day workshop titled “What is essential for Virtual Reality systems to meet military human performance goals” (The Hague NL, April 13 – 15, 2000).

1.2 Purpose and scope of the workshop

The focus of the current workshop is on:

- the functional requirements of potential virtual reality military applications,
- the state-of-the-art and projected capabilities of virtual reality technologies, and
- future research requirements and directions for military applications.

In the light of this focus the following military application domains were considered:

- training;
- robotics;
- remote operations;
- command and control.

Within each of these domains VR requirements, capabilities and R&D issues were considered with respect to the following aspects:

- visual, haptic, auditory and motion feedback;
- navigation interfaces;

- scenario generation;
- modelling software and rendering hardware.

1.3 Program workshop

The workshop took place over three days with the following structure:

Thursday, 13 April 2000

chair:	Pascal Hue (FR) & Thomas Alexander (GE)
focus:	Functional requirements for military VR applications
key-note speaker	Prof. Dr. Roy Kalawsky (Advanced Virtual Reality Centre, Loughborough University, UK)

8 speakers:

- A Virtual Environment for Naval Flight Deck Operations Training (Dr V.V.S.S. Sastry, UK);
- Debriefing for Pilots in VR (Major B. I. Johanson, DE);
- Probing in VR (Mr. L. Todeschini, FR);
- Acquiring Real Worlds Special Skills in VR (Dr B. G. Witmer, USA);
- Training System STINGER Simulator (Dipl.-Ing. M. Reichert, GE);
- Performance Measurements in VR (Dr J. Patrey, USA);
- Appropriate Use of VR to Minimise Motion Sickness (Dr W. Bles, NL);
- Human Computer Interactions in Shared VE (Prof. Dr B. Loftin, USA).

Friday, 14 April 2000

chair:	Elizabeth Henderson (UK) & Lisbeth Rasmussen (DE)
focus:	Available VR techniques now and in the near future
keynote speaker:	Prof. Dr Grigore Burdea (Human-Machine Interface Laboratory, Rutgers University, Piscataway, NJ, USA)

8 speakers:

- Simulating Haptic Information with Haptic Illusions in VR (Mr. A. Lecuyer, FR)
- Tactile Displays (Dr J. van Erp, NL);
- Virtual Cockpit Simulation for Pilot Training (Dipl.-Ing. K.-U. Doerr, GE);
- Ergonomic Investigations for VR (Dr C. Meyer, GE);
- UAV Operations Using VR (Dr Ing. L. van Breda, NL);
- Productive Application of VR (Dr A. Roessler, GE);
- The Dangerous Virtual Building, an Example of the Use of VR for Training in Safety Procedures (Dr M. Lozano, SP);
- Vizualization of Geographic Data in VR (Dipl.-Ing. T. Alexander, GE).

Saturday, 15 April 2000

chair:	Trond Myhrer (NO) & Steve Goldberg (USA)
focus:	Missing VR capability and future research

4 speakers:

- Influence on the representation of Spatial Information Acquired in Virtual Environments (Prof. Dr E. Heineken, NE);

- Development of Virtual Auditory Interface (LCDR Dr R. D. Shilling, USA);
- Educational Conditions for Successful Training with Virtual Reality Technologies (Dr A. von Baeyer, GE);
- Entertainment Technology and VR (Dr M. R. Macedonia, USA).

Panel discussion (HFM021 panel, audience participation).

1.4 Attendees

Attendees (total 43) of the workshop had various nationalities and backgrounds:

Country	Total #	military	industry	academia/ civil res. inst.
Bulgaria	1	1		1
Denmark	2	1	1	
France	5	2		3
Germany	10	1	1	8
Georgia	2			2
Netherlands	7			7
Norway	1	1		
Spain	1			1
Sweden	1	1		
United Kingdom	4	2		2
USA	9	7		2

2 TECHNICAL-SCIENTIFIC SITUATION OF MILITARY VR APPLICATIONS

2.1 Introduction

Functional requirements

In his keynote lecture Prof. Roy Kalawsky (Advanced Virtual Reality Centre, Loughborough University) provided a ‘steppingstone’ for the session on functional requirements. He pointed out that the two most crucial characteristics of Virtual Reality (VR) are the experience of ‘being’ in the simulated world (immersion) and the interaction through sensors (acting). Therefore VR is an essentially ‘man-in-the-loop’ simulation. He pointed out that VR, is not new. The first notions of VR date back to 1956 (Stanton). In fact, a functional decomposition of existing VR systems shows a continuum of levels of immersion starting with non-immersive systems such (e.g. wearable or desktop displays, ‘joystick driven’, with low-level interactions) to fully immersive systems (e.g. head-slaved displays, natural interactions, haptic feedback, etc.). As a simulation tool, the perceived military benefits of VR are mostly related to system effectiveness and not to weapon effectiveness. (not sure what this last sentence means)
The most important message brought by Kalawsky is that the human factor should be central in application development:

- functional requirements (image quality, display scene motion, content development) must be driven by the end application (based on task analysis);
- military requirements must be performance driven in which human capability is a major factor; and
- applications must be evaluated thoroughly with respect to human task performance.

Note that human task performance is governed by the environment itself, personal capabilities, individual motivation as well as the situation under which the task is carried out.

An approach involving this sort of task analysis was illustrated by Reichert (Training system Stinger simulator). Interestingly, TNO developed and spoke about a similar application of VR to training STINGER operators at the previous RSG-28 workshop in Orlando. Based on defining of training goals and tasks the functional requirements of the simulator can be specified (scenario's, interactions, minimum visual resolution, etc.). Although Stinger simulators have been built and used for training, systematic measurements of user performance and transfer of training have not yet been carried out. Sastry presented a study on the use of Virtual Environments for helicopter deck landing training in which training transfer was explicitly measured. Preliminary conclusions show that immersive VE can be used to train visual motor skills. For training procedural knowledge, simpler training devices can be used, although it may be more cost-effective to have a single system for both types of training. Todeschini (Probing in VE) also measured transfer but in the context of VR training for mine clearance tasks. He demonstrated the importance of force-feedback in such applications.

Underlying tasks such as Stinger launching, flight deck operations and mine probing are more basic human skills and abilities. Military personnel need to be able to navigate, orient themselves and interact with the Virtual Environment. The second day of the workshop brought together human factors researchers working in these areas. Witmer presented research on how well people can learn their way around in a virtual world. More specifically: how can we support the learning of routes and configurations VR by adding visual and aural cues in VR? This research yielded concrete guidelines for designing a VR in which task performance depends substantially on spatial orientation and navigation (see Witmer: Acquiring real world skills in VR). Bles (Appropriate use of VE to minimise motion sickness) presented work on modelling the functioning of human equilibrium sensors (vestibular system) and showed which types of motion do and do not cause motion sickness.

Another basic functional requirement, particularly in multi-user applications, is the role of (non-verbal) human communication (human representations and behaviour). One extremely promising development (mentioned by Kalawsky) is the use of avatars (synthetic visual human representations) in Virtual Environments. Avatars can be driven by instrumented humans immersed in the VE or computer generated with Artificial Intelligence programs driving their behaviour. Loftin (Human Computer Interactions in shared VE) presented working which software agents drive human representations (avatars) in distributed shared VRs aimed at training soldiers in peace-keeping operations. A major benefits of avatar-populated VRs is the reduction in the need for

having all the players in a scenario be represented by human being(not everyone needs to be in the loop) which gives the flexibility of just-in-time training.
Important research questions are the required fidelity of the physical appearance of avatars and how to model, generate and validate useful behaviour? Can avatars really be surrogate team-members or coaches? When should they be reactive, when pro-active? When should they behave rule-based, when stochastic?

Besides human factors functional requirements VR design is driven by operational requirements (e.g. mobility, weight, flexibility) and economic requirements (e.g. cost and return on investment). These issues were addressed by Johanson who demonstrated the potential of a mission debriefing system for the Danish Airforce.

Designing VRs that integrate task analysis, functional requirements and technology concessions can be an iterative process that is time consuming. New approaches are needed. Patrey (Performance measurements in VR) presented an alternative approach to traditional cognitive task analysis (rapid interactive design) in order to speed up application developments (see also Loftin). This raised the question of how to interpret performance measurements to adjust system parameters without explicitly modelling the task: should we take a ‘neural net’ approach? (I don’t know what this sentence means)

Available techniques

The state of the art of VR technology was presented by Burdea in his keynote presentation (Available VR now and in the near future). He gave an excellent overview of developments in:
- computing engines (e.g. Intergraph Pentium III based system nowadays match the computing power of SGI Infinite Reality systems);
- tracking devices (e.g. inertial/ultrasonic trackers);
- personal displays (e.g. light weight, high resolution HMO's);
- large volume displays (e.g. CAVE); and
- haptic displays (e.g. haptic gloves, haptic floors).

Burdea observed that:
- VR technologies are getting cheaper (displays, sensors, engines). This brings VR research within reach of research organisations with low budgets for capital equipment;
- consequently, we see a stronger involvement of experimental psychology in identifying the limitations of human performance in VR applications and in validation studies.

The lower costs of VR technology has l allowed for the development of VR applications, for example Doerr’s slow-cost virtual cockpit. Also, 3D worktable technologies are allowing battlefield information to be presented realistically(Alexander). This is an example of how command and control tasks can be supported with VR-tools.

By making use of human information processing capabilities (e.g. illusions in multimodal perception), perceptions can be created without the need of high-tech display devices**.** For example, LeCuyer (Simulating haptic Information with haptic illusions in VE) showed that the perceived haptic amplitude of a spring is substantially affected by the visual representation of the spring and vice versa. It should be noted that a qualitative use of such illusions is feasible but a quantitative use requires individual calibrations. In fact, Roessler (Productive application of VR) stated that a 6D (six degrees of freedom) user interface was even closer to reality without forces than with (for the application discussed).

VR can be used to transfer information through sensory modalities that are not normally used for that purpose. An example was shown by Van Erp (Tactile Displays) who used the skin to sense spatial direction (which is usually sensed by our eyes or ears). By using arrays of tactile micro-vibrators on the skin the position or direction of objects in the 3D space around the user can be presented without putting a load on the visual or auditory modalities. Using such VR techniques, the presentation of information can be prioritised and re-routed depending on the situation in which tasks have to be performed (time pressure, workload, context).

VR technologies can save costs. For example, remotely flying an unmanned aerial vehicle (UAV) requires high-band width (and thus high-cost) video connections. Low bandwidth connections generally yield a limited field of view, low update frequencies and latencies. Together these effects substantially decrease operator performance (overshoots, missing targets). Van Veen (replacing Van Breda) presented a series of studies showing that UAVs can be successfully flown even at a low bandwidth by using VR as an interface technology. This is done by embedding the camera image in a virtual world (augmented reality) in which the visual feedback following an operator’s action (e.g. rotating the camera) is anticipated and thus overshoots are reduced. Therefore, search and fly performance are increased. Van Veen showed that using VR as an interface can be equivalent to a band width increase by a factor of 400.

Auditory displays have not received the attention of visual and haptic displays in the VR community. This is surprising given the fact that 3D sound displays are highly developed, low cost and highly effective. Shilling (development of Virtual Auditory Interface), showed the application of 3D sound in cockpits can reduce the time to complete an attack in some situations by almost 40%! It is also surprising that most VR modelling tools ignore 3D audio (Kalawsky). Also not available are models to represent the multiple modalities of human information processing.

Conclusion

To conclude we can say that current research shows a need for:
- natural interaction devices in VE (usable, intuitive metaphors, multimodal);
- alternatives for cognitive task analysis in designing VE applications (rapid application development);
- team interaction in VE (models of human and object behaviour and social processes);
- design guidelines.

2.2 Bottlenecks and opportunities

During the workshop discussions forty participants from military organisations, academia and industry put forward their opinions on the biggest bottlenecks and opportunities in the development of military VR applications. Listed below are the applications mentioned by workshop participants that represented the greatest opportunities for VR technology(the number of times an item was mentioned is given between brackets):
- (17) training;
- (6) planning, mission rehearsal and debriefing;
- (3) integration of simulation and operation (real missions, augmented with VR);
- (1) remotely operated systems;
- (1) product design.

The most important bottlenecks mentioned are:
- (15) Most VR developments are not user driven: insufficient involvement of the users, insufficient co-operation between designers, users and human factors people, lack of natural interfaces, not enough attention for motion sickness and display quality;
- (6) Lack of standardisation. This makes it hard to integrate systems and software tools;
- (5) Not enough budget;
- (4) Not enough knowledge/imagination: behavioural models for people as well as objects, scenario generation methods, no ‘out of the box’ ideas (people use VR to do the same things they always did).

Obviously the sample of attendees of this workshop perceive training as the greatest opportunity for VR applications which is reflected by the number of presentations on this subject (Sastry: VR training of flight deck operations; Johanson: low-cost mission debriefing system; Todeschini: VR training of mine-clearance). VR for training can

reduce cost and risk of casualties and improve flexibility and performance monitoring. At the same time a number of factors seem to frustrate successful applications in this field (lack of attention for human factors, lack of standardisation, lack of money, and a lack of knowledge on how to model human and object behaviour.
The consequences of a lack of attention to human factors were mentioned by Prof. Kalawsky in his keynote presentation:

- poor user interfaces and crude input devices;
- poor multi-sensor integration (inconsistent visual, auditory and proprioceptive feedback);
- poor facilities for team interactions (poor visual human representations and communication tools);
- a parameterisation of immersion (as an assessment metric) has been fruitless so far.

2.3 Recommended actions

The way to overcome the bottlenecks mentioned above could be:

- identification of a killer application in the field of training (focus);
- involve human factors experts in the development of this application;
- develop VR design guidelines (see Kalawsky).

demonstrate convincingly the value of VR to the military (budgets).
To this purpose the military should develop a vision on the use of VR technology and more clearly specify their needs. Industry should work on standardisation and should substantially bring human factors into their development process. Academia and research institutes should co-ordinate and accelerate their long-term research efforts to focus on natural interfaces (innovative metaphors) and on how to model human and object behaviour. In the short term academia should focus on human factors metrics and metrics for team performance (cognition, communication), and a standard evaluation methodology (Kalawsky).

Specific suggestions made during the workshop which could contribute to solving the bottlenecks are:

- Establish an open NATO specialist group to:
  - identify killer applications;
  - identify a target list of user requirements and technologies for investment by the military;
  - foster development of behaviourally realistic intelligent agents and models;
  - bringing together interdisciplinary groups and create common vocabulary on shared problems;
  - create a research network and identify new funding sources;
  - share software libraries and create a central depository of devices and modules; and
  - open the non-classified publication of results to other organisations.

In general, better co-ordination between military organisations, industry and academia is necessary in order to identify gaps in current knowledge and co-ordinate research.
The enthusiasm of the workshop attendees and the evident willingness to share ideas and to discuss their findings provide a promising base for such co-operation. At the end of the workshop the attendees had formulated a unanimous request for follow-up meetings to work on, exchange and monitor progress on the above mentioned points. This could be implemented in the form of an annual workshop on military applications as a satellite of a major conference on Virtual Reality (e.g. VR2000, organised by Burdea).

3. CONCLUSIONS AND RECOMMENDATIONS

What is essential for Virtual Reality Systems to meet Military Human Performance Goals?

Answers to this question centre on the three focuses of the workshop -- functional requirements, state of the art, and future directions. Day 1 of the workshop detailed the military requirements from which we derive performance goals. Prof. Kalawsky told us that the environment, personal capabilities, individual motivation and the overall situation govern human performance. Thus, the first partial answer is *Military Human Performance Goals include interacting within the training environment in the same way we will interact in the real environment -- train like we fight*. Day 2, state of the art, spoke to the techniques and technology available in the marketplace, both commercial and military. Speakers set the baseline for the VR systems that exist today. Participants discussed in small groups the issue of what might be the bottlenecks and roadblocks to maturing the technology, so that the military potential would be fulfilled as well or better than industry applications. In the highly competitive world of entertainment and automobiles, non-productive techniques don't last.
Simply stated, the second partial answer is that *baseline applications are solid in the automotive industry and entertainment industry, and military applications are beginning to emerge and be evaluated*. Day 3 looked to the future of VR. Work in considering VR for teaching mental representations of knowledge, e.g. spatial knowledge, for enhancing the sense of presence (3D audio effects), and for educational intervention techniques, such as enhancing quality, quantity and retention of skills. The third partial answer to the workshop title question, then, is *the successful military application of VR depends first upon multi-disciplinary implementation teams of scientists, engineers, practitioners, and users, and secondly upon continued advancement of technology toward increased fidelity to the real world*. So, in the ensuing years following Sutherland's 1970 "Scientific American" article that introduced the phrase virtual reality, we have seen literally thousands of projects emerge and we are beginning to see return on investment. However, we need continued investment and synergy for military goals to be met.

3.1 Future Work

Future work is divided into discussions of near and far term work. The intent here is to give the policy makers of NATO an idea of what is emerging shortly versus what will need continued investment. In the near term we can expect emergence of the following

- Human performance measures derived in VR more quickly than from the real world;
- Solutions to side effects that allow prolonged exposures to VR;
- Usability guidelines for VR at the level of detail that we now have for GUI;
- New metaphors linking past knowledge to new concepts being taught;
- Behavioural models of human stress, emotion, fatigue, anxiety and other human traits;
- Intelligent tutors, agents, and behavioural models that enhance the cognitive challenges of training;
- Wearable computers that mix reality and virtual reality to produce superior performance;
- Networking for collaborative work from design to implementation to decision making;
- Visualisation techniques that consolidate vast data into comprehensible information;
- Smaller, faster, cheaper technology from industry, and NOT necessarily meeting military needs.

In the longer term, the military needs a much closer synergy with academia and industry. The trends of reduced personnel, reduced budgets, more accountability, increased demand for return on investments and the expanding military role in operations-other-than-war, will continue to strain the resources and limit the financial influence of the military upon VR training technology development. In effect, the longer term strategy does not yet exist that would give NATO members the science fiction sense-of-presence of the holodeck or that of the movie, "The Matrix," where the effect was so real that humans couldn't tell the difference between what was reality and what was not. That longer-term strategy should exploit the near-term emerging technologies and attempt to influence the direction of longer-term investments by industry and academia.

3.2 Future Meetings

The simplest answer to the question addressed by the workshop is *continued involvement by NATO members in the application of VR technologies to meeting military requirements*. Since VR is an integration of technologies to include modelling, simulation, graphics, haptics and audio, and human factors considerations a multidisciplinary approach is needed. Likewise, NATO will want a central focus of military applications of this very critical training technology. Rapidly reconfigurable, low cost, highly effective training environments don't exist. Pick any two of those criteria and the third becomes unachievable. Yet the promise of VR is the possibility of all three for military training. Such a potential seems well worth continued investment, influence and involvement by NATO member countries.

In addition to M&S, educational applications may also provide an avenue of approach. Many countries are investing in Internet capabilities for their citizens. The network will soon be as comprehensive as the telephone and television. VR immersion technologies and distance learning principles coupled with broadband Internet distribution would eliminate the need for military capital investment and allow cost effective delivery of

training materials. Thus, collaboration of military agencies and civilian educational agencies may be a synergy for combined resources, and long term technology development strategy that would provide the critical mass necessary to influence industry and academia toward training needs. NATO member nations could factor this strategy into some of the thinking about operations-other-than-war.

Medical applications present a further avenue. The Human Computer Interface is currently unacceptable for complex systems. Keyboard, joystick and mouse instruments will give way to EEG, voice, haptic and eye interfaces as technology moves toward human-centred design and network-centric warfare. Already, medical applications of VR appear viable for training surgery and diagnostic procedures. Similarities of human functions need further exploration. For example, France reported mine detection training to be very similar to training for medical personnel to insert a needle. Continued analysis for functional similarities between and across disciplines such as medical applications would give additional leverage to military operations training techniques. The work in metaphor development for VR is one step in this direction. Also, continued understanding of internal human communication mechanisms may provide better Human Computer Interfaces than currently exist.

One conclusion is clear. There is no obvious strategy, no clear consensus and no simple combination of techniques to achieve military performance goals using VR. Three possible strategies are presented, above.

Appendix A: Distribution

- RTA Director for approval of publication in proceedings;
- HFM-021 members;
- Workshop Attendees.

Keynote Address: What is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?

Roy S. Kalawsky
Advanced VR Research Centre[1]
Loughborough University
Loughborough, Leics, LE11 3TU
UK

Summary
The origins of virtual reality can be traced back to the late 1970s and early 1980s with the pioneering research in military crewstation design involving electronic cockpits during. Many of the enabling technologies have even earlier origins than this. More recent developments in the commercial world have resulted in remarkable improvements to some of the limitations of the early generation systems. As the cost of the technology falls and the computational performance increases there is a growing need to ensure that a VR system is optimised for both the user and the tasks to be carried out. Unless the complexities of the associated user interface are understood and carefully controlled there is a high risk that future VR systems will be extremely difficult to use and may be completely ineffective. Sadly, the thrust of most research groups is focussed towards improving the technology without attention to human factors. It is tempting to try and relate the user's performance in the real world with that achieved in a virtual environment. However, before this can be done it is important to establish whether or not it is valid to make such comparisons. This paper focuses on the need to develop a reliable methodology to address the complex human factors issues.

Perceived Military Benefits
The development of military based applications of virtual reality is driven by the following perceived benefits:

- Improved cost-effectiveness — through process integration across the life cycle
- Improved quality of decision making
- Enable teamwork within MOD and with other agencies
- Better understanding of defence issues e.g. Human aspects of warfare
- Focus on system effectiveness rather than weapon performance.

VR is a human centred interface
Even though VR has been evolving for many years we still do not have a reliable or robust definition for VR. The early definitions are themselves becoming outdated as new interaction techniques are developed. To provide a common reference for the term VR the NATO Research Study Group (HFM-021/RSG-28 produced the two definitions below:

Virtual reality is the experience of being in a synthetic environment and the perceiving and interacting through sensors and effectors, actively and passively, with it and the objects in it, as if they were real.

Virtual reality technology allows the user to perceive and experience sensory contact and interact dynamically with such contact in any or all modalities.

Overall, these definitions are appropriate for completely synthetic environments but do not fully address the definition of an augmented reality system involving both synthetic and real environments. The synthetic environment is used to augment or 'fill-in' information in the real environment. From a military perspective the augmented reality system is a very important class of VR system because the technology allows additional information (such as tactical data) to be overlaid onto the real environment. A good example of the use of a synthetic environment is a computer generated terrain displays that is overlaid onto the real world through head up or head mounted displays.

The reason why it is important to amend the original NATO definition of VR to include augmented reality is to acknowledge the different human factors issues an AR system provides. The following definition should be appended to the NATO definition:

Virtual reality can be used to augment the real world and compensate for missing sensory information or to enhance the real world in a way that does not normally exist.

How Best to Describe a VR System: A User Centred AR Taxonomy
The basis for a human factors review of a complex user interface is a functional description of the important interface characteristics. In order to develop a functional description or taxonomy for a VR system it is important to define the scope of the system being investigated. Rather than take a technological perspective, it is much more useful to take a user centred view. This has the advantage of ensuring that human factors issues are properly represented. This approach has already been used with success (Kalawsky, 1996), and Figure 1,

[1] For correspondence with author: r.s.kalawsky@lboro.ac.uk, tel. +44 (0)1509 223 047, fax +44 (0)1509 223 940.
See also: http://sgi-hursk.lboro.ac.uk/~avrrc/index.html.

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

below, outlines the sensory modalities and system interfaces of a generic AR system.

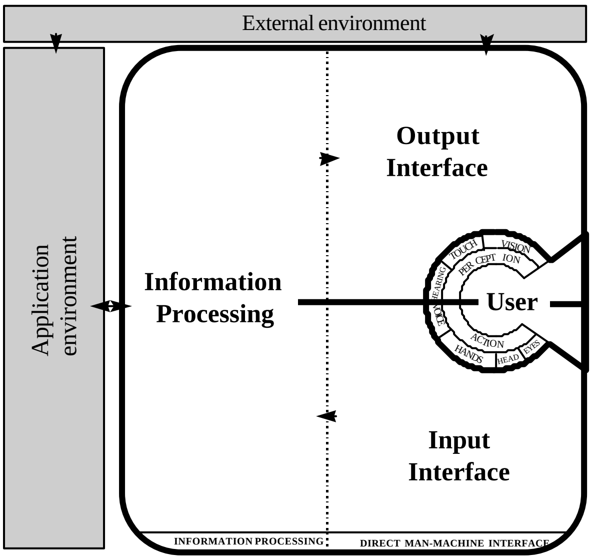


Figure 1: Top Level Functional Decomposition/Taxonomy of any Generic Human-Computer Interface

The functional decomposition of VR, illustrated above, has many uses:

- It can be used to describe pictorially any VR system and allows easy comparison with other systems;
- The diagram can be populated with current technology or future technology capabilities
- It is possible to used linked cells in the decomposition to describe associated human factors issues (either requirements or known problems).

Deconstructing the Framework

When dealing with the components of the taxonomy in more detail it becomes apparent that not only are various technologies catered for but also inherent in the diagram are descriptions of the human factors issues and underlying processes of human factors integration.

Direct human-machine interface: This refers to the user and the functional interface/devices used to experience and control a VR environment.

User: The user is defined in terms of sensory/perceptual processes (e.g. visual, auditory, kinaesthetic, tactile and olfactory) as well as actions that can be initiated by (voice, hands, head and eyes). For completeness, the olfactory sense is included and although current VR systems do not exploit this sense, research in this area is starting to emerge.

Output interface**:** This refers to techniques that can be used to provide information to the human perceptual system.

Input interface: This refers to the means by which human initiated actions can be converted into appropriate information for use in the environment.

Information processing: This is where data is processed for delivery by the output interfaces. Data from the input interface are processed and used to control the environment. This section also has links to the application environment that governs what is actually undertaken by the overall VR system.

Application environment: This is the simulation software that dictates what the VR system will do in accordance with external input from the environment or whatever is initiated by the user. There is a close relationship between this and the information processing section. Example application environments include training systems, flight simulations, molecular modelling, assembly plants etc., and it is feasible for the application environment to be networked to other local or remote applications.

External environment: This represents the real (physical) world that may be linked to the VR system. For example, in a medical application it is feasible to overlay a virtual image onto a patient via an optical system. To achieve accurate registration of the real and virtual environments it is important to provide a link between the display technology, application environment and information processing sections.

The functional decomposition can be broken down into lower levels of detail as required and partitioned as shown in Table 1.

Functional Category		
A. Information Processing		
Output	Cognitive Agents	A1
	Data Management	A2
	Control-display Coordination	A3
	Data storage and Recording	A4
	Image Generation	A5
	Tactile Stimulus Generation	A6
	Kinaesthetic Stimulus	A7
	Auditory Signal Generation	A8
Input	Speech Processing	A9
	Switch Processing	A10
	Virtual Hand Controller	A11
	Head Sensor Processing	A12
	Eye Sensor Processing	A13
	Physiological Sensor Processing	A14
B. Direct Human-Machine Interface		
Output	Image Display	B1
	Tactile Feedback	B2
	Kinaesthetic Feedback	
	Audio Production	B3
Input	Speech Transduction	B4
	Hand Operated Controls	B5
	Head Sensing	B6
	Eye Sensing	B7
	Physiological Sensing	B8
	External Environment Viewing	B9
	Visual Defect Correction	B10

Table 1: Functional Category Pointer to Descriptive Tables.

Figure 2 shows how the functional decomposition is applied to an augmented synthetic environment (augmented reality) system. Each cell in the functional decomposition corresponds to a descriptive pointer

where more detailed information is stored. A description of the sort of data that is stored in the table can be found in (Kalawsky, 1996) and details factors such as, current technical specifications, future technical requirements and human performance implications.

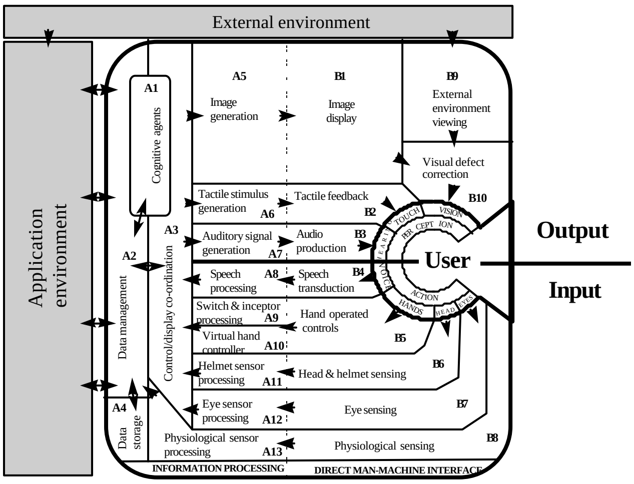


Figure 2: Detailed Functional Decomposition of an AR Interface

The numbers used in the diagram act as pointers into a series of descriptive tables (Table 2) that are used to describe the technical specification of the enabling technologies (current, predicted, or even novel future concepts).

Functional Category	Potential Use	Likely Technique	Limitations
B. Direct Human-Machine Interface			
B1. Visual Information Display			
Desk top displays	High res., non immersive display	CRT, LCD, Large screen CRT, Plasma Projection display	No correction for curved screens
Head coupled displays	360° field of regard Full immersion, high res., low lag display	CRT, LCD colour shutter,	Single user mode, Low - medium res, Light transmission Field of view
B9. External Environment Viewing	Integration of real/virtual environments Overlay of virtual display onto real environment Augmentation of real environment	Optical, electronic mixing, Chroma key techniques	Registration between real and virtual environments

Table 2: Example Extract from Low Level Cross Referenced Data

Classes of VR System

From a technical perspective it is convenient to categorise a VR system according to the degree of immersion it provides. In this context, immersion refers to the extent that the user is enveloped in a virtual environment and is related to the technology employed. Three degrees of immersion have been defined as:

- Fully Immersive
- Semi-immersive
- Non-immersive

Fully Immersive VR Systems

Fully immersive VR systems are characterised by being able to completely envelop the user with a synthetic environment wherever they are. It is tempting to think only in terms of the visual channel but other modalities such as auditory perception are equally valid. However, in the majority of applications, the visual channel will be the most dominant and the auditory channel will be used to augment the visual channel.

Head Mounted Displays

The development of head mounted displays can be traced as far back as the early 1950s. They were the first display technology to deliver a fully immersive experience and since the early 1990s there have been many different designs for the head mounted display. These tend to fall into two categories — non see-through and see-through. As these terms imply, the non see-through head mounted display does not allow the user to see any part of the real world. Conversely, the see-through head mounted display makes it possible for the real world to be overlaid with computer generated graphics generated by the head-mounted display.

Although it may seem that the see-through and non see-through head mounted displays are similar they actually present very different human factors problems that must be considered in the context of the application and operating environment.

Non see-through head mounted displays

The non see-through head mounted display typically comprises two display devices (typically CRT or LCD) and a set of optics to magnify and position the image a fixed distance from the user. The image can be presented anything from a few meters to optical infinity (beyond 250 meters). The exact distance is usually a design feature of the head mounted display and is fixed by the display manufacturer. The image plane distance is extremely important and is a function of the application.

There are a range of non see-through head mounted displays available and these literally come in many different configurations.

Technical Description

There are many different configurations for non see-through head mounted displays and it would not be practical to review them all in this paper. Refer to (Kalawsky, 1993b) for a more detailed account. Figure 3 shows a simplified non see-through head mounted display with a simple magnifying lens. The distance of the virtual image from the user is governed by the distance of the image source (typically a CRT or LCD) from the focal point of the magnifier lens. If the image source is located at the focal point then the virtual image is located at optical infinity. In practice the head mounted display manufacturers tend to position the

virtual image much closer than this, typically 1–20m away. There does not seem to be a particular preference so the image distance can vary from one design of head mounted display to another. For most tasks, this does not make too much difference except perhaps where it is important to visualise large scale building structures at a scale of 1:1. In this case, the virtual image should be placed as far away as possible. However, in the case of a see-through head mounted display, image distance is extremely important.

Figure 3: Non see-through head mounted display

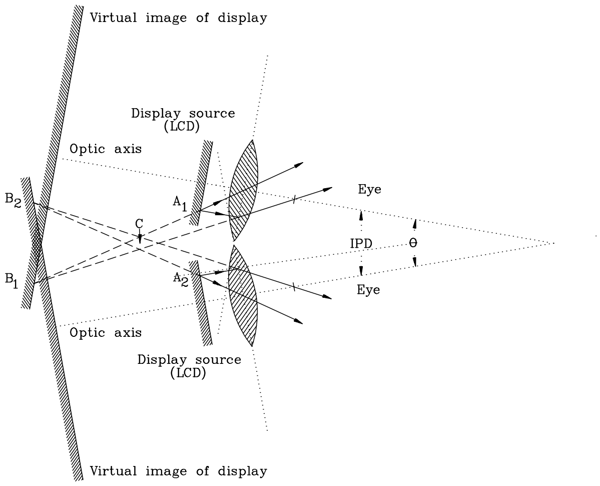


Figure 4: Relationship between Image Source and Virtual Image of a Non See-through HMD

Head mounted display technology still lags behind the requirements of most applications. Notably, the display resolution is still far too low to be of practical value. It should be stressed that display resolution must not be considered in isolation. An equally important and related parameter is the field of view of the optical system. If an application calls for a narrow horizontal field of view (for example, 40°) then a display resolution of 1280×1024 might be adequate. However, if the required horizontal field of view is in excess of 140° then this would probably be inadequate. There are so many other trade-offs that have to be considered that it is no wonder an off the shelf head mounted display is unable to meet a particular requirement. As a comparison, in the military sector head mounted displays are specially designed for each application. Unlike commercial applications where one display is intended to fit all applications.

User Issues

There is no doubt, head mounted displays are unpopular with potential end users. Apart from the above problems comfort is a major factor. Current off the shelf head mounted displays are still too bulky for many people and after short periods of use people report discomfort in areas of neck strain, eye strain, claustrophobia and nausea. The general health and safety issues of head mounted display are now being understood and the enabling technology is being improved gradually.

One of the biggest drawbacks of a head mounted display system is that it provides a single person experience whereas the current trend in VR is for group or multi-user interaction. As soon as the user puts on the head mounted display they are isolated from the real world.

Despite the technical difficulties associated with head mounted displays progress is being made with development of higher resolution display sources. Whether or not these developments make it possible to reconsider the use of non see-through head mounted displays remains to be seen.

Despite these concerns the applications where non see-through head mounted displays can be considered include:

- Large scale architectural visualisation where large screen systems would be impractical.
- Maintenance training
- Research involving phobias where it is important to isolate the user from the real world.

See-through head mounted displays

The more exciting though technically more challenging head mounted displays are the see-through systems. Interestingly, the very first head mounted displays were based on optical systems that overlaid display information over the real world. These systems were later developed to become an important system for fighter pilots. A number of very sophisticated systems have been developed. Commercially available see-through head mounted displays are now available and are based on cheaper versions of the military systems. The term augmented reality (AR) is frequently used to refer to see-through head mounted display systems. Computer enhancement of the external environment offers distinct advantages over virtual reality by not only potentially avoiding the need for complex modelling of people and the environment, but also by providing an

anchor in reality that should reduce the likelihood of nausea being induced. Instead of replacing the real environment with one that is completely artificial, a number of early researchers (e.g. (Sutherland, 1965); (Furness, 1986); (Knowlton, 1977); (Krueger, 1985) (Kalawsky, 1992, Kalawsky, 1993b)) have used computers to augment the real environment. Augmented reality systems offer the potential to allow user the to actively carry out tasks involving real world objects rather than being confined to an artificial environment such as is the case for virtual reality based systems. Figure 5 shows a modern commercially available see-through head mounted display.

Figure 5: Sony Glasstron Augmented Reality Head Mounted Display

Technical Description

There are two main ways in which the real world can be augmented by a graphical overlay:

a. A see-through head-mounted display can be employed enabling the user to see the real environment through part-silvered mirrors that also reflect a visually superimposed graphic image into the user's eyes. The optical system relies on a partially reflecting semi-transparent surface providing an integration of the real world with information generated by an electronic display such as a cathode ray tube (CRT). The external environment is generally viewed through the combiner plate and the image from the CRT is reflected by the combiner plate. Refer to Figure 6.

b. A conventional VR head mounted display can be used to provide a non-see-through augmented reality display in which the user sees a video image of reality combined with luminance or chroma-keyed graphics (Kalawsky, 1991). An AR system based on electronic overlay relies on a video mixing system taking video from a television camera viewing the real world scene and superimposing it with a video signal from a computer graphics system.

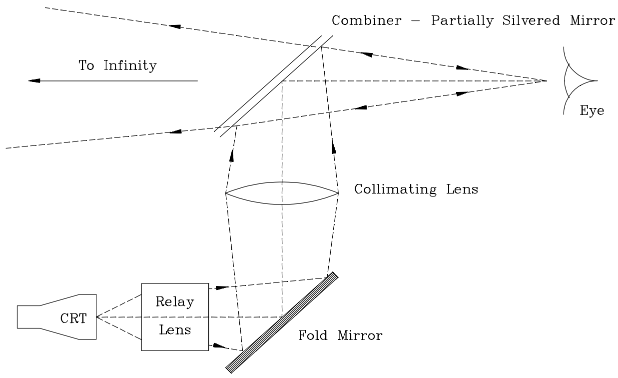


Figure 6: Basic Optical System for See-through HMD

The majority of head mounted displays for commercial applications have been predominantly non see-through. The reason for this may be the great difficulty that is experienced when an attempt is made to register the virtual display with objects in the real world. Any misregistration that arises from calibration errors, lags in the graphics or head tracking system etc. is immediately apparent to the user (Kalawsky 1992a; Kalawsky 1998). For the tasks suggested for see-through systems (maintenance, design etc.) the misregistration has proved to be very problematical.

User Issues

Although AR concepts have been around since the 1950's the technology and its application is still in its infancy. This in the main, has been due to technological limitations of synthesising real and virtual images in the same visual field, and fundamental problems of image registration and collimation. In recent years, AR systems have become more sophisticated and offer particular advantages over VR concerning some of the human factors issues that arise. For example, in the case of AR, orientation cues are still available to the user from the visual scene in the real world. Users are therefore unlikely to experience the feelings of vertigo and sickness that can be brought about by traditional VR systems (Caudell, 1994). However, AR configurations produce unique issues of their own.

Research into the human factors issues surrounding the use of AR systems is very limited and few formal guidelines exist for any application of AR technology.

Irrespective of which technique is used to provide the electronic display overlay there are several technological factors that must be considered. These include:

- Image plane position of the virtual image
- Transparency (or rather the transmissivity/reflectivity) of the combiner assembly.
- Registration accuracy of the electronic image with respect to the external environment.

Each of these factors will have an influence on how and what information is displayed to the user.

Image plane position: All virtual display devices produce an image at a particular position from the eye. The position of the virtual image can be as far away as optical infinity and is controlled by the position of the image source with respect to the collimating lens. For example, if the image source is located on the focal point of the collimating lens the virtual image is at infinity. Virtual image position is usually set at infinity for aircraft applications but is inappropriate for other applications. Commercial off the shelf head mounted display systems usually fix the virtual image position at some arbitrary distance (e.g. 3 m). For most game applications this distance has not been proven to be critical. However, when these displays are used in conjunction with information derived from the real world (i.e. operated in see-through mode), virtual image position is very important. Unless the virtual image is collimated to be coincident with the information in the real world a misregistration occurs. The net effect on the user is the need to re-accommodate when attention shifts between the information displayed in the real world and that displayed on the head mounted display. Particular care must be taken when using see through HMDs if there is an accommodation/convergence mismatch with the external environment. When operated with such defects it is quite easy for serious eye strain to occur. The long term exposure to such eye strain is not fully understood. Not all see through HMDs suffer from this effect. However, due to possible commercial and legal implications it is not possible to identify the problem-atical see-through HMDs in this report.

Transparency (or rather the transmissivity/reflectivity) of the combiner assembly: The optical design of an AR display will determine what percentage of the real world is transmitted through the display to the user and what percentage of light from the image source is overlaid onto the real world. The nature of the semi-reflecting surface of the optical combiner (beam-splitter) also has an effect. Some devices work by employing a notch filter to maximise the percentage of light overlaid onto the real world. Unfortunately, this has the effect of removing certain spectral components from the real world. Obviously, the impact of this depends upon the application.

Registration accuracy of the electronic image with respect to the external environment: Registration accuracy is very important in head tracked augmented reality display systems because mismatches between the information presented in a virtual overlay compared with the real world may affect user performance. There are two types of misregistration error. The first is caused as a result of a static misalignment of the virtual image with the real world. If this error is present, it can usually be calibrated out. However, the second type of misregistra-tion error is a temporal misalignment caused by delays in the computer and tracking system. It is possible to create an AR system with almost zero static misregistration errors but as soon as movement occurs computational delays introduce misalignment.

Advanced embedded training systems offer the potential to train operators in their real working environments rather than spending time being trained elsewhere. In the future, such systems may provide on-line feedback to operators, perhaps tailored to different levels of operator expertise, offering dynamic and flexible alternatives to conventional training facilities (Zachary, 1997).

Semi-Immersive VR Systems

Semi-immersive VR systems represent a very exciting class of system because they overcome many serious problems associated with head mounted displays. Distinct advantages include:- higher-resolution displays, multi-participant experiences, wide angle display. A semi-immersive display does not provide a fully enveloping display image. Depending on the display technology used a field of regard of up to 270° can be obtained. Field of regard refers to the extent of the displayed image expressed in angular terms. The term field of view is sometimes used to represent the same thing. However, to be strictly correct, 'field of regard' refers to the instantaneous field of view as perceived by the observer with fixed head position. The field of regard refers to the total display field of view that can be seen by moving the head around. The field of regard is therefore potentially larger than the field of view.

Flat Screen Systems

Though not normally considered by some to represent a semi-immersive display system, it is feasible to refer to flat projection screen based systems as a semi-immersive display, provided the field of view is greater than 90° horizontally, Figure 7.

Figure 7: Wide Field of View Flat Screen Projection System (Rear Projection)

Technical Description

The flat screen system can be based on a single or multi-projection display system. The actual arrangement of projectors depends upon the required resolution of the whole system (in horizontal and vertical extent). A single projector can achieve a maximum resolution of about 1600×1200 pixels. (Please note, there are specialised higher resolution projection systems

available but these are likely to be expensive. The more typical resolution is either 1280×1024 or 1024×768. Quite acceptable large screen displays can be produced with these resolutions from either CRT or LCD technologies. If higher resolutions or fields of view are required then it is important to increase the number of projectors. A typical arrangement is shown in Figure 8.

Rear or front projection can be used. The exact choice being dependent on the way users interact with information on the screen. The main problem with front projection is if users need to get close to the screen they can cast a shadow, which obscures displayed information. However, a rear-projected display does not suffer from this problem. Due to the composition of a rear projection system the screen material tends to diffuse the light more than a front projection system and this can lead to an image with poorer contrast.

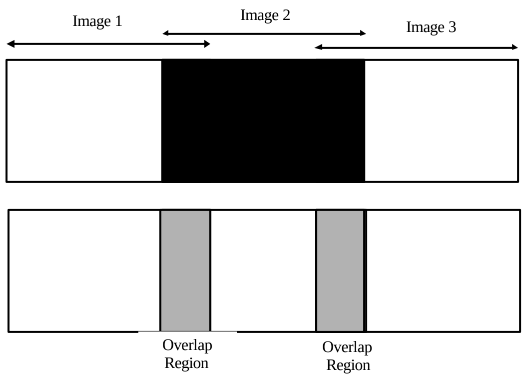


Figure 8: Arrangement of a Flat Screen Multiple Projection Display

Whichever projection method is used it is very difficult to match the display output from different projectors precisely because of the errors (optical distortions) present in all projection lenses. The magnitude of the error increases the further you move away from the optic axis of the lens system. In order to combine two images together it is necessary to overlap the image of one projector with another projector by a few degrees. Special electronics units (known as an edge-blenders) are used to match the image edges together. The more sophisticated edge blenders enable an accurate colour balance to be achieved between the overlapping projected images. In the past, there have been attempts to perform the edge blending in the graphics system but this adds to the computational complexity of the system. Unfortunately, this affects overall system performance and edge blend deficiencies become very noticeable in dynamic display imagery.

A further issue is the physical placement of the projectors. In other semi-immersive systems it is desirable to position the projectors close to each other but in very large flat screen systems this is not feasible because the projector has to compensate for large keystone errors. This is illustrated in Figure 9.

User Issues

The flat screen display systems present fewer user issues provided the user does not rely on display peripheral cues too much. If a single projector is used then the need for frequent and sometimes difficult alignment between projectors is avoided. If a CRT based projector is used instead of a LCD based display then the user must be prepared to re-converge the three CRTs occasionally to maintain the system at optimum performance.

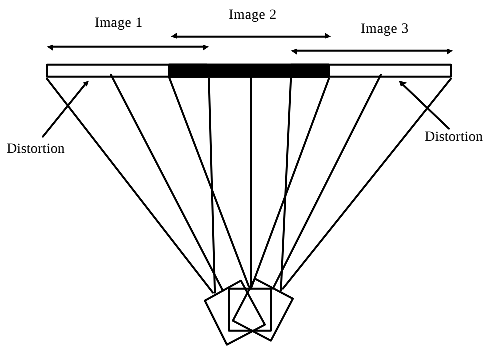


Figure 9: Arrangement of Projectors to achieve Lower Levels of Distortion

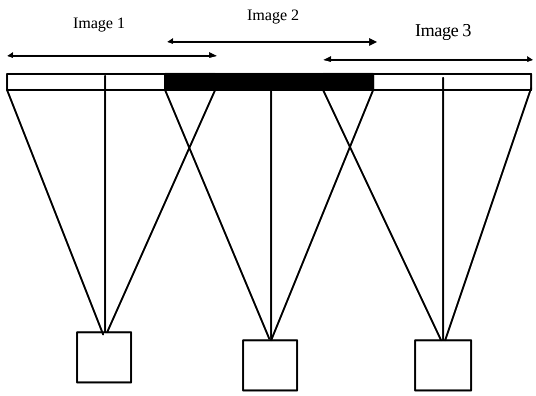


Figure 10: Arrangement of Projectors for Flat Screen

Over time, the three CRTs will slowly drift out of alignment and the edges of objects displayed on the screen will have a coloured ghost like image drawn around them in one or more of the primary colours. This indicates that the CRTs are out of alignment. Obviously, in a multiple projector system the user must be prepared to re-converge each projector and to align each projector with respect to each other. Changes in the thermal environment in the laboratory frequently account for this misalignment. It is worthwhile considering the use of a temperature controlled environment for multiple projector systems as this will certainly help reduce the number of times projector alignment is carried out.

The flat screen system can be used for a wide range of applications. The flat screen semi-immersive display is without doubt a very cost effective way of creating a compelling display environment. In its simplest form, a single projection system can be used requiring only a single graphics system. As more and more projectors are used the complexity of the graphics system and the requirement for an edge blending system increases the overall cost.

Immersive Workstations

The immersive workstation is a term used to cover a number of small flat screen projection systems that provide a powerful visualisation capability. These systems are very simple in concept and can be bought for relatively low cost.

Figure 11: Immersive Workstation

Technical Description

The configuration of the immersive workstations is very simple. It relies on a high brightness projector and a rear projection screen. The screen is orientated at an angle rather than being placed vertically in front of the user, Figure 12. Some immersive workstations allow the whole table to be rotated through 90° from horizontal to a vertical position. Figure 12 shows a very simple diagram of a typical Immersive Workstation. The fold mirror is used to simply keep the size of the unit to a more manageable size. Without the fold mirror the required distance between projector and screen to create a large image would be impractical for many installations.

The stereo display is produced in a frame sequential manner whereby alternate left and right eye images are presented by the projector. In order to see a stable stereo image the user must wear special glasses that shutter the left and right eyes in synchronism with the projection of left and right images. A small infra-red transmitter is used to send a signal to the stereo shutter glasses so that the left/right eye shutter can be correctly synchronised.

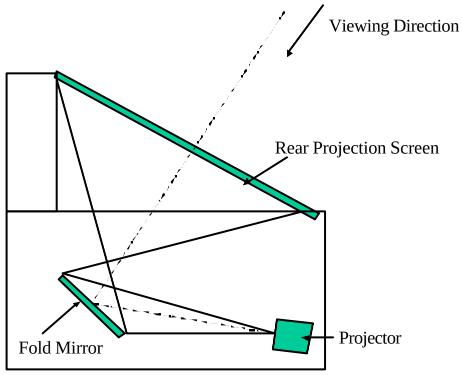


Figure 12: Schematic Diagram of an Immersive Workstation

User Issues

Immersive workstations are very useful tools for visualising 3D objects in stereo mode. Users should be aware that they offer a fairly restrictive viewing/ operating area due to the nature of the stereo display system. It is not possible to provide a correctly computed view for more than one person in head tracked mode. If other people wear the stereo shutter glasses, they will see a stereo image provided they are in front of the display. However, the image will not be geometrically correct for their viewing perspective. This may not be an issue for many applications.

The immersive workstation is a very versatile device and can be very effective for small group interaction (1–4 people). However, if head tracking is employed then the person linked to the head tracking system will be the only person able to see geometrically correct stereo images.

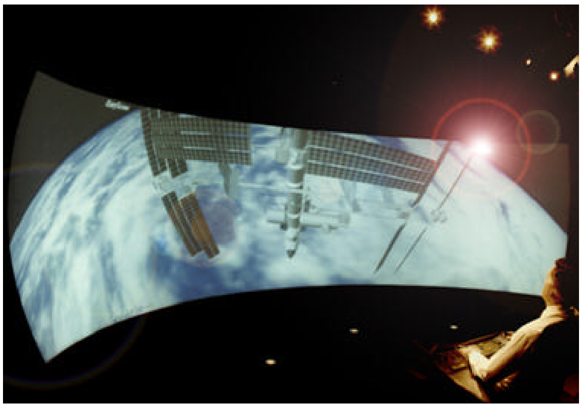


Figure 13: Silicon Graphics Reality Centre – Theale (Reading)

Reality Centre Systems

A further development of the multiple projector flat screen system has lead to the evolution of a curved screen system, Figure 13. This is extremely similar to the flight simulator used in the aerospace industry albeit without the cockpit. Instead of a cockpit, a group of users are situated at the focal point of the curved screen.

The main characteristic of the curved screen system is that the user perceives a greater field of regard and the edge distortion effects noticeable in a flat screen system are almost eliminated. Single or multiple projectors are used, the number being a function of the required display resolution in horizontal and vertical extent.

Silicon Graphics coined the term Reality Centre™ to cover this type of system.

Technical Description

The basic principle of a Reality Centre lies with the arrangement of the projection system which is situated at the focal point of a curved screen (Figure 14) whose horizontal extent is anything from 90–180° horizontally. To achieve the wider fields of regard it is usually necessary to employ multiple projection systems and overlap their screen edges. A video edge blender is then used to blend the edges of different projectors together in a way that make the overall image appear as a single uniform image.

User Issues

The curved screen Reality Centre is a very convenient tool for many applications. The curved screen currently rules out non CRT projectors (e.g. LCD) since it is not possible to incorporate distortion into the image to compensate for the curved screen. As higher resolution projectors become available it will be possible to replace multiple projector configurations with a single projection system. This will greatly facilitate system maintenance and remove the need for edge blending systems. Setting up and maintaining a single projector can be quite time consuming if continuous peak performance is required. If multiple projectors are used it is necessary to achieve alignment of each projector against a projected reference image. Over time, the CRTs used in the projectors will age at different rates and between channels. The better edge blending technology will permit some degree of colour/matching between adjacent projectors as each CRT ages. However, if the system is used frequently then it might be advisable to rotate the projectors around or be prepared to swap out the CRTs on a more regular interval.

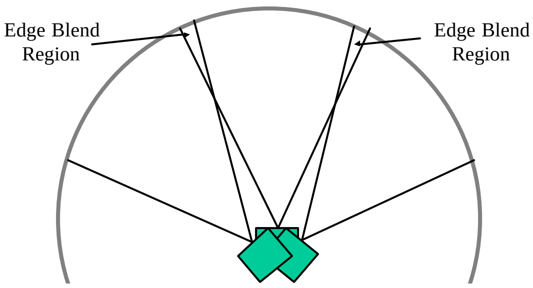


Figure 14: Diagrammatic Representation of Reality Centre Screen System

In order to reduce the amount of re-calibration it is advisable to maintain the temperature of the room where the Reality Centre screen and projection is installed. Temperature drift is one of the main causes of the image going out of alignment.

Reality Centres can be used in all sorts of application though the educational value has yet to be determined. One of the particular strengths of a Reality Centre along with flat screen and Vision Dome systems is that they are ideal for groups of people. The cost of owning such a facility can be very high but alternative ownership schemes such as leasing might prove to be extremely attractive. A number of organisations have established Reality Centres as a commercial centre where they hope to sell time on the facility to external organisations.

CAVES

A very interesting development has taken place with the flat screen projection system. Instead of employing a single flat screen, several screens are used at right angles to each other Figure 15. By arranging for each screen to be orthogonal with respect to each other, it is possible to create a room called a CAVE, whose walls are formed from rear projector screens. The CAVE™ is a multi-person, room-sized, high-resolution, 3D video and audio environment. The CAVE was developed at EVL http://www.evl.uic.edu) and is available commercially through Pyramid Systems Inc. There have been many earlier examples of such screen-based systems involving rear-projected displays. Many of these have originated in the Aerospace industry. In 1987 the author saw a very early example at Wright-Patterson Airforce base. From 1990 British Aerospace used a multi-faceted display system for its cockpit research programme. These early generation systems were not known as CAVES but had all the properties of today's CAVE systems.

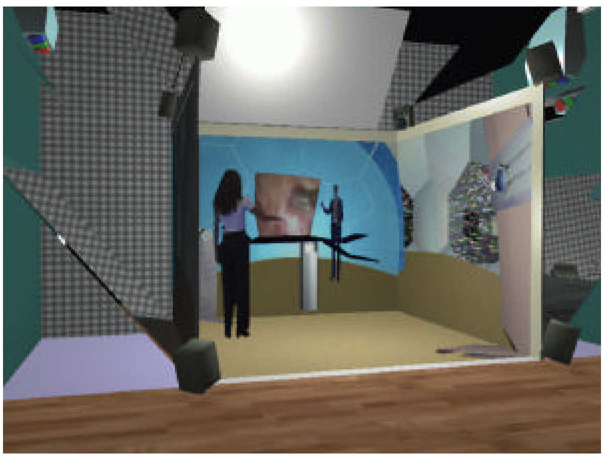


Figure 15: CAVE Display

Technical Description

There are many variations of the CAVE concept but they all rely on the principle of the user being surrounded by three or more orthogonally arranged rear projection screens. Figure 16 shows a top view of a three-sided CAVE system. The three projectors are situated outside

the inner projection viewing area. In practice, the size of the image on the walls of the CAVE is a function of the lens used and the projection throw distance. It is possible to extend the number of sides to a CAVE up to the maximum (i.e. six) by carefully positioning additionally projectors around the CAVE.

It is desirable to use as large a CAVE as possible. This in turn requires a large space in which to site the CAVE and associated projection equipment. The space requirements of a six sided CAVE should not be under-estimated. In order to reduce the amount of space required to realise a CAVE, mirrors can be used to increase the effective throw distance as shown in Figure 17.

Obviously, the cost of a CAVE system increases as a function of the number of sides is employed. Each side will require its own dedicated graphics channel and projection system.

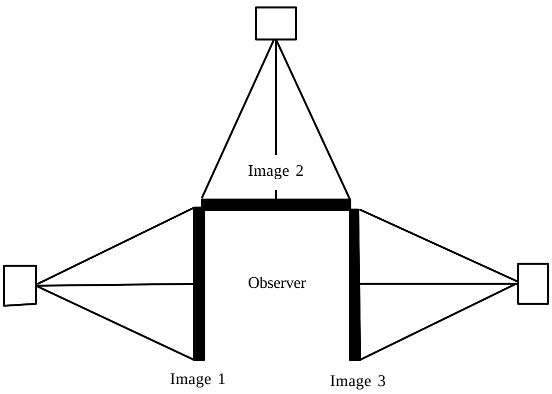


Figure 16: Three Sided CAVE Configuration

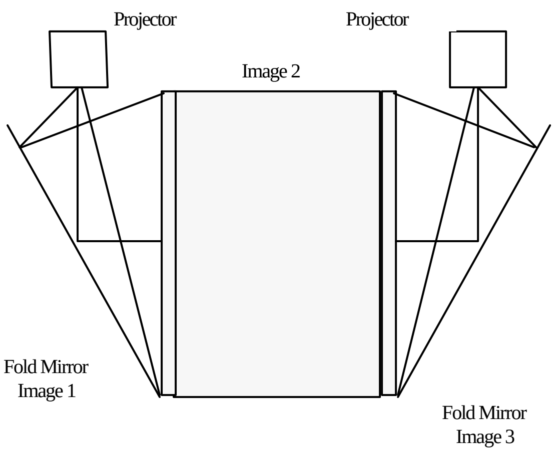


Figure 17: Arrangement of Fold Mirrors to Increase Effective Throw Distance and Reduce CAVE Space Requirements (Note Front View)

The blending of edges in a CAVE becomes very challenging because of the abrupt angular changes that occur at the junction of the sides of a CAVE. Dedicated edge blending technology exists that will match the geometry of the corresponding points of one CAVE side with another. The edge blending will always be a compromise because of the viewing geometry with respect to the angle of the sides of the CAVE.

A requirement of all projection display systems is the reduction of veiling glare caused by reflection of light from surrounding areas from the projection surface. In a CAVE the light from one side will inevitably be reflected from the opposite side and cause a significant reduction in contrast. Many CAVE users frequently complain about the poor contrast displays.

User Issues

A six-sided CAVE can provide a total immersive experience for one user. The display presented on each wall of the CAVE can be a stereo image in which case the user perceives a display with depth. Care must be taken to carefully calibrate a CAVE system otherwise the user can experience nausea effects. Due to the nature of a head-tracked CAVE, if other users are present, they will obtain a distorted image because they will not be at the same viewpoint of the person being tracked. The technology does not exist yet whereby multiple users can be tracked in the CAVE so that each gets a correct view.

One of the main problems present in head mounted display systems is the accommodation/vergence effect. The accommodation/vergence system is tightly coupled in the human perceptual system. When the user fixates on an object the accommodation response is partly driven by the eye vergence. Any errors in the object's distance perceived by the vergence system and the accommodation response will cause eye strain for the user. The head mounted display produces an image at a fixed focal plane. A CAVE system can similarly present accommodation problems because for a given viewing position the observer's eyes have to re-accommodate if the object under view is displayed on two or more display surfaces. Even if the CAVE system can be calibrated, it is not possible to compensate for the different accommodation required as an object is viewed. It is possible that people may develop nausea or motion sickness that was symptomatic of head mounted display systems.

The veiling glare problem briefly described above is often the source of complaints of poor contrast by users of the CAVE system. If the background scene can be kept quite dark (not always possible with some virtual environments) then the veiling glare can be kept to a minimum.

If the user in a CAVE rotates or rolls the image, it is possible to induce sufficient visual cues to interfere with the user's balance so that they fall over. This is

particularly the case for a six-sided CAVE where it is possible to loose sense of true horizontal. This phenomenon is exploited in some fairground ride systems and people do actually fall over. Some people are more susceptible to these effects than others.
A CAVE system is very expensive so this limits their use to applications where the cost can be justified. The following application domains are evaluating CAVES:

- Automotive
- Architectural
- Art
- Oil and gas sector

Vision Dome Systems
The Vision Dome concept is not a new idea. It is based on the astronomical planetarium which instead of projecting a film-based image onto a spherical surface, the film projector is replaced with a CRT based projector. The Vision Dome theoretically can present a full 360° field of regard image but in practice the full field of regard is seldom used. Figure 18 shows the Loughborough University Vision Dome.

Figure 18: BT/Loughborough University Vision Dome

Technical Description
The Vision Dome comprises a hemispherical projection screen that is driven by a single projector located at the focal point of the screen. Figure 19 shows a typical Vision Dome installation.

The necessity for a single projection system at the focal point of the hemispherical screen places a requirement for a very high-resolution projection system. The optical system has been designed to cover the whole hemispherical surface and the projection system remains as the limiting factor.

The screen used in the Loughborough University 5m Vision Dome is quite interesting in that it is maintained by air pressure. An internal aluminium structure comprises two layers (one layer is the screen). Air is drawn out from between the two layers and this forces the screen to take on a perfect hemispherical shape.

Figure 19: 5m Vision Dome Schematic

The single projector means that a single graphics pipe is required. However, the spherical nature of the screen requires real-time distortion correction to be applied to each image before it can be displayed in the vision dome. Fortunately, many high performance graphics systems can cope with the increased computational load.

Portable versions of the Vision Dome exist and employ similar air pressure maintained screen systems. Apart from being smaller they do not require the large external support structure. This means that erection of the assembly takes a matter of a few hours instead of several days.

User Issues
The Vision Dome does not present an accommodation conflict as in a CAVE system to the user because the image plane is maintained at a consistent distance from the user across the whole field of regard.

One of the definite advantages of the Vision Dome compared with a CAVE is the use of a single projection system. Therefore, the need for matching different projectors and display surfaces is eliminated. However, there is a need to employ a special graphics library that replaces standard graphics calls with modified versions that take into account the required distortion correction for the spherical dome surface. Modified graphics libraries are available for most platforms including NT based PCs. Since only one projector is required this means that the projector must be extremely high resolution to cover the whole field of view. Unfortunately, requires a more expensive projection system but means that a single pipe graphics system can be used with a significant reduction in cost.

Effective 3D user input devices are still required in common with all other VR systems.

Non-Immersive VR Systems
Desk-top Systems

It is considered unnecessary to review in any detail the use of non-immersive desktop VR system since these are based on conventional display monitors. It is possible to drive many display monitors in frame sequential stereo mode and achieve the benefits of an immersive workstation.

User Issues

Without a doubt, a desktop CRT monitor will give the best image in terms of the following:

- Resolution
- Contrast
- Clarity
- Colour gamut

For some very critical users, there is simply no alternative technology. Even modern LCD monitors fail to compare with the highest quality CRT display.

Portable VR — Wearable Computing

VR technology is normally associated with large fixed installations but advances in wearable computing technology have made it possible to produce mobile VR systems. There are a number of military initiatives around the world looking at how the future solider would be deployed with technology. One such programme is known as Dismounted Infantryman. Loughborough University are addressing some of the complex human factors issues of wearable computers, Figure 20 shows the second generation system.

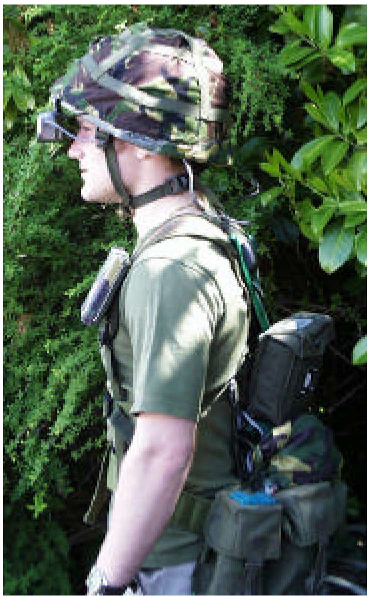


Figure 20: Loughborough University - In Field Computer Mark 2

The main operational feature of wearable computing systems are:

- Hands free interaction
- Contextually aware
- Always on and assisting the operator

The use of high powered portable computing devices presents a whole new series of human factors issues. These will not be discussed in this paper because of space constraints.

Human Factors Issues of VR Systems

The functional requirement of a VR system is driven by the end application and must take into account human capability (physical & cognitive).

Sensory Conflict!: Real World versus Virtual Environment

In order to deal with the complex area of human performance and effectiveness (the goal of human factors) it is important to note the distinction between real and synthetic environments. The most appropriate way of doing this is from a user's perspective. The user will generally experience sensory conflict in the synthetic environment, even though they may not be immediately aware of the effects. It is easy to recognise the effect of sensory conflict when ones compares a real world experience such as a roller coaster ride with a video of the same experience. The sensory inputs to the rider on the roller coaster will be through, vision, sound, smell and proprioception. The rider will also experience a wide range of rich sensory cues such as air rushing over the face, the sound of the roller coaster, other riders screaming and shouting, sense of vibrations, extreme inertial forces, accelerating, turning and descending and the intense emotions of fear and excitement. In contrast to this, a video of a roller coaster ride provides typically two sensory inputs, vision and sound. Not only are the number of sensory inputs limited they also tend to have a lower fidelity than in the real world. Additional effects are also present such as temporal lags introduced by the video system. All these inherent features reduce visual fidelity of the experience. It is also possible that some sensory cues may even be contradictory.

The 'Perceptual Sense of Being' in a Virtual Environment — Presence

An important differentiating characteristic of VR systems compared with other human-computer interfaces are their ability to create a sense of 'being-in' the computer generated environment. Other forms of media such as film and TV are also known to induce a sense of 'being-in' the environment. Some VR practitioners have tended to use the term presence to describe this effect (Sheridan, 1992), (Heeter, 1992), (Kalawsky, 1993a), (Zelzter, 1994), (Hendrix and Barfield, 1996). This means that people who are engaged in the virtual environment feel as though they are actually part of the virtual environment.

In order to understand what it means to be present in a virtual environment it is necessary to understand what characteristics of the real world enable us achieve a sense of presence. A good example of real world experience is a roller coaster ride. The sensory inputs to the rider on the roller coaster will be through vision,

sound, smell and proprioception. A roller coaster rider will experience air rushing over their face as well as the sound of the roller coaster and other riders screaming and shouting. The sense of vibrations and extreme inertial forces when accelerating, turning and descending, etc will be very real. It is obvious that most people will also experience intense emotions involving fear and excitement. A video of a roller coaster ride there are only two sensory inputs, vision and sound. Both of these sensory inputs will be much less real than the real world. Stereoscopic depth cues will be absent from visual and auditory information. Temporal lags introduced by the video system will further reduce the visual fidelity of the experience. Some sensory cues may even be contradictory. The body will feel comfortable in a normal seating posture but the visual cues (with reference to the horizon) will be indicating that the body is anything but stable. If the roller coaster is experienced in an IMAX cinema the reaction of others will have an effect. Some people report that they can suppress the sense of presence by weakening or strengthening their awareness. Upon receipt of sensory information, some people can fill in gaps to create a better or enhanced sense of what is happening. For example, people who have previously experienced a roller coaster ride would experience a different state of awareness than someone who had never experienced the ride. This implies that previous experience may affect the sense of presence.

Intersensory Interactions

Traditionally, sensory modalities have been investigated in isolation from another. It has been suggested by (Sherrington, 1920) that all parts of the nervous system are connected together and no part is capable of reaction without affecting or being affected by other parts. This means that examination of part of the system will inevitably lead to an incomplete understanding of the perceptual experience. Intersensory interaction relates to the perception of an event when measured in terms of one sensory modality which is changed in some way by the concurrent stimulation of one or more other sensory modalities. Given the nature of the human sensory system there is great diversity in the intersensory interactions that can be experienced and this adds to the difficulty in understanding what is happening.

Spatial Location

There are at least four sensory modalities that are capable of providing spatial information to the human being. These are visual, auditory, tactile and proprioception. The visual sensory modality is the most spatially acute of the spatial modalities with a resolution acuity of about 1 min of arc. In contrast, the ability to spatialise a 1kHz tone placed in front of the participant's head at varying angular distances from the median plane of the head gives a minimum angle of about 1°. Tactile acuity is a very difficult thing to define because it depends on what part of the body is being stimulated. The tongue has a two-point threshold of about 1mm.

Orientation

There are four sensory modalities that support the perception of orientation: visual, tactile, proprioception and vestibular sense. Proprioception is a very powerful mechanism for conveying a sense of body orientation though it has been shown that with time the body can adapt to unusual positions and this can lead to false orientation cues being perceived. The visual and vestibular senses are extremely accurate in conveying a sense of orientation of gravitational direction. This is one of the reasons why the perceptual system can make serious errors in orientation judgement if one of these two sensory modalities is missing or conflicting with the other. Misperception of the body and the gravitational direction vector can cause a shift in auditory localisation cues (Graybiel and Niven, 1951). This phenomenon is known as the audiogravic illusion.

Egocentric localisation

Egocentric localisation is the ability of the human to perceive the direction and distance of objects relative to the observer. Egocentric localisation is achieved by the visual, auditory, tactile and proprioception sensory modalities. It is usual for several of these modalities to act together to give an accurate sense of localisation.

In the real world it is common for several of sensory modalities to receive simultaneous stimulation in a way that reinforces a common multi-modal perception. However, in the virtual environment system it is possible that one or more of the sensory modalities will receive incorrect stimulation due one of the sensory channels not being provided. This phenomenon is sometimes referred to as intersensory bias.

Whilst we can examine sensory interaction and relate this to specific human capability, the term presence has defied all attempts to define it in a quantifiable manner. There is clearly a coupling between the senses and the phenomenon of presence (Gilkey, 1995). Gilkey has examined the level of presence experienced by suddenly deafened adults. These deaf adults frequently complain of a sense of unconnectedness with their surroundings, which supports the view that auditory cues are important for establishing a sense of presence. It is unfortunate that many people make the mistake of assuming that the most important cue in a virtual environment is the visual modality. Even if we concentrated entirely on the visual channel there would still be sufficient auditory cues around in the real world (including self generated auditory noise such as breathing) to limit the sense of sensory depravation that is reported by suddenly deaf people. Interestingly it has been reported by (Gillingham, 1992) that acoustic isolation and lack of auditory cues may account for spatial disorientation.

The term immersion is also sometimes used erroneously to describe the experience of presence. The term immersion in fact refers to the extent of peripheral display imagery. If the display presents a full 360°

information space then we are dealing with a fully immersive system. However, if the extent of the display is less than this we have a semi-immersive system. The term non-immersive is usually reserved for desk-top VR systems. To avoid confusion it is best to associate immersion with the technology characteristics of the display. Unfortunately, these terms are not interchangeable and refer to quite different things. Presence is essentially a cognitive or perceptual parameter whilst immersion essentially refers to the physical extent of the sensory information and is a function of the enabling technology.

Perceptual Conflicts — Phantom Illusions

Gibson (Gibson, 1986) has mentioned the notion of co-perception of one's own movement, in other words awareness of locomotion. The visual system "is kinaesthetic in that it registers movements of the body just as the muscle-joint-skin system and the inner ear". In the real world, the visual system perceives information about the environment and one's own self in that environment. Our whole perceptual system behaves in this manner and processes many reinforcing cues from the environment. It is better to think of these cues as reinforcing since they are all contributory rather than some current views that suggest these cues provide a degree of redundancy. Visual kinesis is a powerful perceptual process as evidenced by a wide-angle panoramic projection screen. It is quite easy to produce very convincing and compelling visual cues that give the participant a sense of self-locomotion. The visual experience can appear to untrained people as a very vivid illusion of reality even though the participant is anchored to the floor. A similar illusion can occur whilst sitting on a stationary train in a station and an adjacent train pulls away. Sometimes, you become convinced that you are moving and the other train is stationary. It comes as a surprise when you discover that you are in fact stationary. What is very interesting with these experiments is the way visual cues can override cues from the vestibular system. Although it is tempting to isolate a particular sensory modality when trying to explain perceptual phenomena it is problematical.

Someone who is completely blind would argue that they can 'see' the environment through auditory, haptic and kinaesthetic cues. Indeed, when deprived of the visual channel you soon become aware how extremely important the other modalities are. By allowing, the person to move their head and move within the environment, proprioception fills in much of the information that would normally be provided by the visual channel. The presence of all sensory modalities removes some of the ambiguities that can occur with a reduced set of sensory inputs.

Great care must be taken not to infer that everyone behaves in the same manner. Some people are far more sensitive and can compensate for conflicting sensory cues than others.

In the majority of experiments conducted in presence, the experimenters do not address the issue of sensory conflict. It is quite possible that our real-world experiences which are based on a full set of sensory cues do not readily map onto our sense of presence in a sensory deprived computer generated environment. Not only is our sensory system deprived of certain perceptual cues there may be sensory conflicts, which arise from issues such as lags or temporal anomalies in our system.

Our experiences or priori knowledge of real world systems can greatly influence our internal representation of a sense of being present in an environment. For example, test pilots are used to dealing with tasks in a fixed based (no motion cue) simulator and transfer the experience to the real world. However, it has been established that most combat pilots perform better in simulated missions compared with real battle situations. Obviously, risks are much easier to take in a simulator than in the real world. As a converse argument combat pilots sometimes make different decisions when under combat stress due to a different level of adrenaline. Unless people are carefully trained (and it is very difficult to determine if this can actually be done) then there is a great danger that the subjective evaluation techniques may not be sensitive to the same parameters for each of the experimental participants.

A computer-generated environment can affect the participants experiences in a very profound way by allowing events or situations to be experienced that cannot be achieved in the real world. For instance, it is easy to transport someone to a different temporal domain where events can be slowed down or speeded up compared to real time. In these situations, it is not practical to try and map this onto a real world experience. Consequently, researchers should be very careful when using terms such as low and high presence.

A crude but repeatable measure for presence would be to count the number of sensory inputs that are missing from the virtual environment compared with the real environment (Sheridan, 1992). Unfortunately, even this approach is flawed because each sensory modality does not contribute equally to the sense of presence. It is also likely that individual contributions will change over a period. For example, it is well known that people can become desensitised to certain stimuli.

Real World Versus Virtual Environment

There is considerable merit in being able to compare performance in the real world against performance in a virtual environment, especially if the virtual environment is mimicking the real world in some way. This means that metrics developed for the real world case can be deployed in the virtual environment. However, this presumes human performance is the same in real and virtual environments. This factor is very important for training applications where a virtual environment is used to train a particular skill and the skill has to be

transferred into the real world. Skill transfer is a very important factor but equally human behaviour and performance in the virtual environment is very important. Conflicting sensory cues could actually modify the user's performance in the virtual environment in a detrimental or beneficial way. In some cases the ability to present only a subset of real world attributes might actually improve the training process. Pilot training is a good example where basic procedural tasks can be taught without the trainee having to worry about flying the aircraft at the same time. These training systems are known as part task trainers. To avoid many training transfer issues the part task trainer is made as real and representative as possible.

It is possible to extend this idea by investigating the quality of a virtual environment in terms of the tasks to be undertaken. If we gather data on the user's performance in a real environment and the user's performance in a virtual environment then we have some measure of the quality of the virtual interface.

The Quest for Understanding Presence

There has been insufficient research into the causes of presence to be able to discuss them definitely and accurately. "There is no scientific body of data and /or theory delineating the factors that underlie the phenomenon" (Held and Durlach, 1992). Despite this, there is a growing quantity of research that is attempting to derive a single dimension for presence. This research is based on subjective rating techniques. Zelzter proposed a description of virtual reality in his AIP cube (Zelzter, 1994) which sets out to define the components of a synthetic environment in terms of a co-ordinate system giving a measure of the quality of the system across three interacting parameters: autonomy, interaction and presence. Zelzter's cube illustrates the three different axes — defining a co-ordinate system which can be used as a qualitative measure of virtual environments. Autonomy is defined as the ability of the environment to act and react to simulated events. Interaction is the fidelity with which the environment deals with interactions between its participants both human and synthetic. Presence provides a rough (dimensionless) measure of the number and fidelity of available input and output channels. Zelzter associated the position of an application within the cube with task performance. He indicated that while clearly, an evaluation of a virtual reality system in these terms is highly task dependent, every design solution for a virtual environment can be characterised within these bounds.

At first sight, Zelzter's AIP cube looks to be a good way of describing a particular virtual reality system according to where it fits in the cube. However, it is very difficult to characterise a virtual reality system in this way because of the lack of a clear definition for each axis. In particular, the term presence is very difficult to specify in a simple way that would fit the AIP cube. It is tempting to try to classify attributes of a virtual reality system in this way, and then devise a measurement process but there is a serious danger that the real performance controlling factors of a virtual reality system will not be addressed. Moreover, it is not easy to justify the use of a dimensionless performance parameter if it cannot be measured objectively or subjectively against clearly defined metrics.

To begin to understand where and how to evaluate the user's performance when using a virtual environment it is necessary to look further into the unique properties of the system. Traditional empirical human factors based evaluations such as measuring the display resolution of the system are useful but do not necessarily relate too well in terms of overall user performance. For example, it has been shown that performance in a virtual environment is affected if one of the input modalities is removed (Pausch, Shackleford and Proffitt, 1993). This suggests that if we undertake empirical based evaluations we will not be able to draw too many conclusions regarding an integrated interface. In this context, we need to consider the virtual environment system as an entity and thus treat the system as an integrated interface.

Traditional human factors evaluation techniques do not take into account attributes such as presence and greater interactivity. There have been numerous attempts to produce a single metric representing the degree of presence for a virtual environment and then relate this to some measure of human performance (Slater, Usoh M. and Steed A., 1994). Research which investigated the sense of presence (as yet undefined) within virtual environments as a function of visual display parameters (Hendrix and Barfield, 1996). The research indicated that people reported higher levels of presence when head tracking was used and stereoscopic visual cues were employed. An increase in field of view also resulted in a reported increase in the level of presence. In reality, these finding are no surprise since we routinely use the cues in the real-world.

There have been many attempts to define a straightforward definition for presence, largely without success. Some VR practitioners try to define different classes of presence such as ego-presence and object-presence (Hendrix and Barfield, 1996). Indeed, it is tempting to try and derive a simple measure for the amount of presence a particular system is able to provide and then relate this to user task performance. Unfortunately, this approach is flawed because presence is a multi-dimensional parameter that is arguably an umbrella term for many inter-related perceptual and psychological factors. However, it is clear is that presence is a cognitive factor that must be treated differently than other perceptual aspects of a human-computer interface such as brightness or contrast of an image. If presence can correlate usefully with performance and provide the means to achieve effective

communication and control in interface design, (Ellis, 1996).

Issues of Evaluation

A virtual interface is radically different compared to conventional computer interfaces and as such needs quite a different approach to performance evaluation. The user's performance is governed by the environment, personal capabilities, individual motivation, the tasks to be performed and the situation under which those tasks are to be carried out. For instance, if the user is performing two tasks simultaneously then performance on a single task might not be the same as when only one task is being undertaken. Whilst it is possible to perform empirical experiments to predict human performance these do not tend to deal with a complex situation where numerous activities have to be undertaken concurrently. An empirical understanding of human performance is important but what is probably more important is an understanding of the overall user performance. This seems strange when a virtual environment system has the potential for so much variability in the design of the interface.

It is easy to overlook that we are dealing with a multi-sensory interface that can provide auditory, kinaesthetic and visual displays. One point to bear in mind during the evaluation process is that the user will experience fewer sensory cues in a virtual environment than in the real world. This inevitably means that our knowledge, which relates to the real world may only partially fit the case for virtual environments. Indeed, we only need to think of cues such as motion perception to begin to understand the complexity of the problem.

The user of a virtual reality system will generally act inside the environment rather than outside as with other computer based systems. In many ways the flight simulator (one form of virtual environment) has similar attributes. Consequently, a number of interesting human factors challenges will result. For example, if a user is performing a task in a virtual environment it is quite possible that an experimental evaluator (on the outside) will interfere with the performance of the user. Fortunately, this type of problem can be overcome by careful experimental design. In situations where highly realistic virtual environments are being used, the lack of certain real or redundant sensory cues may have a detrimental effect on the user's performance and subjective experience. An equally important issue is one of perceptual conflict where for example, dominant visual cues may conflict with whole body kinaesthetic cues. This has been known to be the cause of many accidents in the aerospace sector where pilots have tended to believe their own proprioceptive senses rather than aircraft instrumentation. In some instances it will not be possible to avoid such complexities, as the enabling technology will be limited. However, this is where an understanding of empirical human performance becomes important. One way of avoiding the difficulties of human performance evaluation is the development of an evaluation framework. The framework could help formalise the whole process and ensure that a consistent approach is taken.

User Interaction Devices

A review of the human factors issues of VR systems would not be complete without a discussion on the user interface. Most VR system users would agree that the user interface is poor and the input devices are relatively crude. Even though there are a variety of different input devices ranging from 3D joysticks to glove like devices none of these are particularly intuitive. This partly comes from the need for an effective 3D interface device and user metaphor. Some tasks will require force feedback and this area is seriously lacking in terms of the maturity of the enabling technology.

If the VR system is to be used to support group interaction then the situation is even more serious because no group interaction devices exist. All the interface devices are severely restricted because of the need for cables and wiring. In the future wireless interface devices will be required.

The VR System as a Collaborative Tool

Arguably one of the best uses for VR is collaborative working at remote locations. It is now perfectly feasible to link two or more VR systems together over computer networks and establish a common virtual environment between the remote users. With such a system it possible for each remote user to interact with the data set and makes changes that are then reflected to all other collaborating users. The justification for collaborative working arises from the complexities of today's projects, which tend to be multidisciplinary and involve teams of people who could work for different organisations. These organisations could be international and one clear benefit of the collaborative link up would be the cost and timesaving compared with travelling to a common destination. The collaborative VR system would not solve the time zone differences but would save considerably on project costs. The collaborative VR system is very different to a video conferencing system because it allows the users to interact with the data being discussed or reviewed.

There is an important issue of scale regarding the collaborative VR system. On one hand there is the potential for collaboration involving hundreds of users whilst on the other it is possible to restrict collaboration to just two–five people. It is very clear to see how chaos or confusion could set in if a large number of people are collaborating together. Figure 21 shows what tends to occur on the user's display. There are many commercially available products that support this type of interaction over the internet. The usefulness of these systems has yet to be proven.

Figure 21: Potential Chaos with Large Numbers of Collaborating Users

There are other available products that enable much more effective interaction (though this is still limited). Figure 22 shows the sort of environment that is provided by Parametric Technologies Corp. (formerly Division) dVMockup.

Figure 22: Collaboration using dVMockup

There are many human factors issues that arise from the use of collaborative VR systems. In any collaborative link up we need to know who we are interacting with. It is not necessarily a good idea to have a cartoon like character representing what we are doing in the virtual environment. Obviously, a question to be addressed is what does the avatar communicate and how should it be represented. This raises social implications such as - Who am I actually dealing with? As shown in Figure 22 it is possible to introduce a perceptual expectation mismatch because users can seem to float up in space and assume all sorts of unusual attitudes. This would not occur in the real world and whether this actually helps in the collaboration process has yet to be determined. Even so it is still reasonable to assume laws of physics hold true since this facilitates our interaction in the environment.

One of the first social issues to be addressed is whether a human form actually needed? However, human forms communicate certain social expectations. In a team, working environment is an avatar misleading for a group/team? One area where an avatar comes into its own is if an autonomous intelligent agent is implemented in the virtual environment. Without some form of intelligent control, the avatar will be crudely driven by simple user gestures. Obviously, full body suits could be produced but these are too immature at the moment and more importantly would anyone want to instrument themselves up in this way.

Human Factors Evaluations in Virtual Environments

Issues of Evaluation

It is important to recognise that a virtual interface is radically different compared to a conventional computer interface. This implies that there are likely to be different human factors issues that need addressing. Moreover, it is unsafe to assume that evaluation methods that work for real-world situations may not necessarily work for synthetic environments.

User's performance is governed by the following:

- Environment
- Personal capabilities
- Individual motivation
- Tasks to be performed
- Situation under which tasks are to be carried out

A key problem with many human factors based evaluations is that they are often. If the user is performing two tasks simultaneously then performance on a single task might not be the same as if only one task was being undertaken.

Better to derive a functional or parametric form representing presence.

Rather than go into specific evaluation issues here the interested reader is recommended to obtain the following paper (Kalawsky, Bee and Nee, 1999).

Conclusions

VR systems have technological limitations which are slowly being overcome. However, our understanding of the human factors issues is seriously lacking. It is not simply a question of understanding how we perform in the real world and simply mapping this onto the virtual environment. As this paper has reported, we have to understand human performance in the context of sensory conflict and misrepresentation. From research under-taken it has been established that our behaviour and performance does not necessarily relate to that which would be achieved in the real world under the same task situations. This need not be a major problem because as we begin to understand human interactions we may be able to exploit more fully the unique properties of the virtual environment. Moreover, as our knowledge increases we should be able to impose more definitive

requirements on the development of the enabling technologies and so ensure that they are more suitable for the task in hand.

References

Caudell, T.P., (1994), Introduction to Augmented Reality, Proceedings of SPIE Proceedings Telemanipulator and Telepresence Technologies.

Ellis, S.R., (1996), Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence", *Presence*, vol. 5, pp.247-259.

Furness, T.A., III., (1986), The Super Cockpit and its human factors challenges, Proceedings of Proceedings of the 30th Annual Meeting of the Human Factors Society; pp.48–52.

Gibson, J.J., (1986),*The Ecological Approach to Visual Perception*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Gilkey, R.H., (1995), The Sense of Presence for the Suddenly Deafened Adult, *Presence*, vol. 4, pp. 357-363.

Gillingham, K.K., (1992), The spatial disorientation problem in the United States Air Force, *Journal of Vestibular Research*, vol. 2, pp. 297-306.

Graybiel, A. and Niven, J.I., (1951), The effect of a change in direction of resultant force on sound localisation: The audiogravic illusion, *Journal of Experimental Psychology*, vol. 42, pp. 227-230.

Heeter, C., (1992), Being There: The subjective experience of presence, *Presence*, vol. 1, pp. 262-271.

Held, R.M. and Durlach, N.I., (1992), Telepresence, *Presence*, vol. 1.

Hendrix, C. and Barfield, W., (1996), Presence with Virtual Environments as a Function of Visual Display Parameters, *Presence*, vol. 5, pp. 274-289.

Kalawsky, R.S., 12 Oct, Integrated Real-Virtual World Display System, Patent Appl.9121707.5,,

Kalawsky, R.S., 24 Apr, Integrated real-world virtual world system, Patent Appl.US 07/959;919, USA,

Kalawsky, R.S., (1993a), The Science and Engineering of Virtual Reality, Proceedings of Virtual Reality International '93,, London.

Kalawsky, R.S., (1993b),*The Science of Virtual Reality and Virtual Environments*: Addison-Wesley-Longman.

Kalawsky, R.S., (1996), Exploiting Virtual Reality Techniques in Education and Training: Technological Issues, ISBN 1356-5370,.

Kalawsky, R.S., Bee, S.T., and Nee, S.P., (1999), Human Factors Evaluation Techniques to Aid Understanding of Virtual Interfaces, *BT Technology Journal*, vol. 17, pp. 128-141.

Knowlton, K.C., (1977), Computer displays optically superimposed on input devices, *Bell System Technical Journa*, vol. **56**, pp. 367-83.

Krueger, M., (1985) VIDEOPLACE: A report from the Artificial Reality Laboratory, in *Leonardo*, vol. **18**.

Pausch, R., Shackleford, M.A., and Proffitt, D., (1993), A User Study Comparing Head-Mounted and Stationary Displays, Proceedings of IEEE Symposium on Research Frontiers In Virtual Reality.

Sheridan, T.B., (1992), Musings on Telepresence and Virtual Presence, *Presence*, vol. 1, pp. 120-25.

Sherrington, C.S., (1920),*Integrative action of the nervous system*. New Haven: Yale University Press.

Slater, M., Usoh M., and Steed A., (1994), Depth of presence in virtual environments, *Presence*, vol. 3, pp. 130-144.

Sutherland, I., (1965), The Ultimate Display, Proceedings of IFIP Congress;pp:506-508.

Zachary, W., Ryder, J., Hicinbothom, J., & Bracken, K., (1997), The Use of Executable Cognitive Models in Simulation-based Intelligent Embedded Training, Proceedings of Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting.

Zelzter, D., (1994), Autonomy, Interaction and Presence, *Presence*, vol. 1, pp. 127-132.

A Virtual Environment for Naval Flight Deck Operations Training

Dr. Venkat V.S.S. Sastry[1]
Mr. Joseph Steel
Applied Mathematics & Operational Research Group
Cranfield University
RMCS Shrivenham
Wilts SN6 8LA, UK

Lt. Edward A. Trott
Defence Procurement Agency
Abbey Wood
Bristol
UK

Abstract
The main aim of this paper is to develop a prototype virtual environment for training Flight Deck Officers with a view to study the types of interactions required in such an environment. The application is ideally suited to exploit techniques based on proprioception, in particular the trainee's arm signals.

1. Introduction

A virtual environment is *a synthetic sensory experience that communicates physical and abstract components to a human operator or participant* (Kalawsky, 1993). Virtual Environments (VE) offer greater potential to enhance the communication between the human and the computer as they offer most intuitive and natural interfaces. They have been exploited in diverse applications ranging from medicine to training soldiers. Their potential is far more evident in training applications (Nemire, 1998) for the enriched interaction styles such environments support. A typical VE synthesises one or more sensory inputs to facilitate a particular user's task. Exploiting proprioception (sensory awareness of parts of the body) enhances interaction in a virtual environment (Mine, 1997). One form of proprioception is the use of body-relative actions called gestures to issue commands to alter the environment. Current research work in this area includes two-handed input (Hand, 1997) and gesture-based interaction (Mapes & Moshell, 1995). The gestures involved in the present application are quite unique, and thus provide an ideal test bed for exploring 3D interactions in VE.

Training Flight Deck Officers (FDO) is an important aspect of Naval operations and currently uses a range of traditional teaching material augmented by instructor-assisted scenario generation (AIR 230 Course). The instructor directly controls the scenario presented to the trainee. While this approach has some strength, we believe that a virtual environment offers much significant benefits and that the application readily lends itself to exploit the natural interaction styles, such as arm signals, that are inherent in the training of Flight Deck Operations (Trott, 1999). At the same time, the application raises several challenges. The main purpose of this article is to present the results of our initial prototype, with a view to enhance the model. The development of the application is by no means complete, and should be treated as an initial investigation.

The paper is organised as follows. A brief motivation is presented in Section 2. The problem we are attempting to address in this paper is described in Section 3, together with typical training scenarios. The details of developing the Virtual Environment are given in Section 4. Some of the points highlighted in developing the prototype are summarised in Section 5.

2. Rationale

Exploiting natural interaction metaphors offered by the application can enhance the current set up for supporting the training of Flight Deck Officers. The main motivation for the present study can be summarised as follows.

- To provide an enhanced training environment for the trainee.
- To allow interactions using natural metaphors that will enhance the experience of a flight deck officer in which he/she will be able to control the environment in response to his/her actions. For example, ask the helicopter to move to the next gate position in response to an arm signal.

3. The Problem

Current Practice

Currently the British Royal Navy Flight Deck Officers (FDO) are trained at RNAS Culdrose, Cornwall, England. Their training makes extensive use of real simulation, that is *real* people using *real* equipment. In this case the real equipment is an actual helicopter, although training is not performed on-board ship.

If the weather conditions restrict aircraft flights or aircraft are unavailable, then the training makes use of a *virtual* simulator. This consists of three large projection screens that display images from three front projectors driven by three networked PCs. The system shows the view as seen from the landing deck of a frigate. The trainee stands in front of the screens and directs the flight of the simulated helicopter using the appropriate signals. The class instructor who is sitting behind the trainee flies

[1] For correspondence with author: sastry@rmcs.cranfield.ac.uk; tel. +44 (0) 1793 785315; fax +44 (0) 1793 784196.

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

the helicopter. The direction of view of the system is fixed and cannot take into account the direction of view of the trainee. The current system also has limited graphics capability and environmental effects such as reduced visibility, fog and variable sea-state are barely implemented.

Although the virtual simulator is available, final examinations must be passed using the real equipment. There are two main reasons for this: the virtual simulator is unable to replicate the feel of being exposed to the prevailing weather conditions nor the feel of the proximity of the helicopter as it lands.

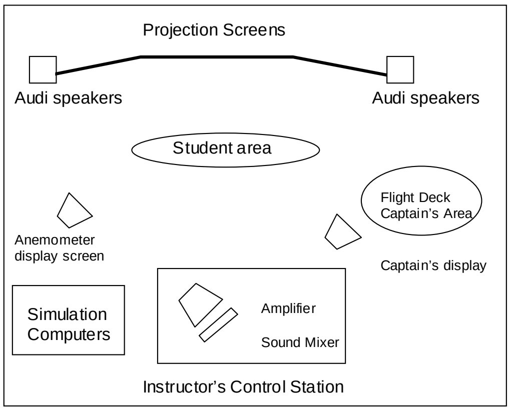


Figure 1: Flight Deck Operations Simulator

Some of the limitations cited above can be addressed by developing a virtual environment, which offers far greater potential. To explore these possibilities, a subset of training scenarios are considered for the initial design, which are explained in the next section.

Example Scenarios

Flight Deck Officer's training consists of a number of scenarios including Landing and Takeoff (varying angles of approach), Rotors Running Refuel, Helicopter In-flight Refuel, Weapons Loading, Personnel Transfer, and Helicopter Shut Down and Start Up using either a Sea King/Lynx. All these scenarios offer a rich variety of 3D interactions that enhance the learning and training experience. For the purposes of this investigation, we have selected two scenarios — Helicopter landing and take off under varying environmental conditions.

The trainee FDO immersed in the virtual environment makes an assessment of the wind speed and direction (not currently implemented) and then signals the aircraft when he is ready to receive it. It is assumed that the aircraft is out of radar-controlled approach and is within the visual range for FDO to take control. On receipt of the signal, the aircraft moves to its next waypoint or gateway. A typical approach of an aircraft to the ship's flight deck is shown in Figure 2.

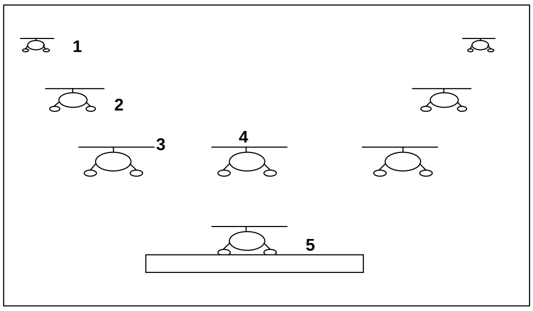


Figure 2: A typical approach of an aircraft. Helicopter begins approach relative angle 165° from ships head (1); Aircraft reaches gate (2); Aircraft alongside flight deck directly of the 'bum' line (3); Traverses across flight deck maintaining its hover (4); aircraft descends to flight deck (5).

4. Development of a Virtual Environment

The virtual environment consists of a visual model of the flight deck and its associated dynamics, a visual model of a helicopter (Sea King) and its associated dynamics, and finally a visual representation of the flight deck officer including body articulation (limited to hands). For the training purposes, few environmental effects such as fog and night time are also included. These are discussed in detail in the following sections.

Flight Deck Officer

A simple model of a mannequin is used to represent the flight deck officer. The body articulation is limited to arms only. Currently there are nearly 56 distinct gestures used by the FDO. Of these, 9 gestures (See Figures 3 and 4) which are directly relevant for launch and recovery operations are chosen for the prototype. An existing model of a man from the Division software library has been modified to facilitate the emulation of arm signals. The animation of the limbs is achieved using the keyframes animation technique where each frame describes a particular state of the object, for example its position and orientation. Each hand signal is stored as a key frame animation sequence and was stored in a separate library.

The FDO is required to carry lighted wands at night in order to make his signals visible to the pilot. Accordingly, our virtual FDO has two lighted wands, which come into effect during night time training.

Helicopter Approach
A 3D model of the Sea King helicopter is modified to provide realistic rotor disc motion and the addition of navigation lights to the helicopter stub wings. This was achieved by mounting a spotlight just in front of the navigation light and setting an appropriate object luminescence property in the object's texture file.

The movement of the helicopter is governed by a series of keyframe sequences in a required direction. The helicopter movement is triggered by an event (such as receive, approach, move away, left, right, up, down, hold, wave off) raised by FDOs hand signal. In response to this signal, the aircraft will move to the next 'gate' position. Once the aircraft has been successfully directed over the deck, a final signal to descend to the deck is given. When the collision between the helicopter and the deck is detected, the helicopter object is parented with the deck object, so that the helicopter moves in accordance with the motion of the flight deck and that of the ship.

Platform Dynamics
The platform consists of the deck, harpoon grid and a model of the FDO standing on the deck. To keep the frame rates to a minimum, a simple animation sequence is created for the platform that emulates the motion of a ship.

Environmental Effects
Environmental effects such as lighting conditions and visibility effects can be easily incorporated into the virtual environment. To effectively light the scene and allow for a number of different lighting conditions five light sources are used. Four of these light sources are used for scenery lighting and the remaining light is used to illuminate the deck. An appropriate texture is applied to the sky to produce an impression of a marginally cloudy day. Fog is emulated using the library function *dvFog* which allows a colour and distance parameter (beyond which the objects are invisible) to be specified. Note that when fog is enabled, the sky is obscured; and affects the intervisibility computations.

Virtual Command Interface
The prototype environment is developed using the Division software dVS/dVISE. Due to the current limitations, the arm signals of the trainee are *simulated* using virtual menus (See Figure 5). A limited number of training manoeuvres is implemented for helicopter landing and takeoff under different environmental conditions. The use of directional sound is also explored with limited success. Note that for a fully functional immersive environment, appropriate hardware, and additional software for gesture analysis should replace the virtual menus mentioned above.

5. Remarks
As the main focus in this study on developing a virtual environment for training, no specific experiments were conducted to evaluate the overall benefit of such a system. However, the exercise has revealed several important factors.
Top most in the list is the need for a software component for gesture interpretation. The computational demands for training purposes are moderate, as the objects in the virtual environment remain fairly static. Selection of items using Virtual menus is not intuitive, and could be enhanced using additional visual cues.

The next stage of the work is an investigation into the recognition of various arm signals using a single tracking device in each hand. Simply knowing where each hand and it's orientation is insufficient. It is expected that knowledge of how each hand has recently moved will be required to determine the relevant signal. For example, arm signals for FDO Up and FDO Down (See Figure 4) trace the same path, but differ in start and finish positions. We envisage that it will not be necessary to have tracking devices at either the elbows or shoulders. The use of a neural network to facilitate this task is expected.

6. Conclusions
The prototype environment for FDO training has highlighted some of the requirements that are essential for a fully immersive tool. There is a clear need to track the position of both head and two arms. While the current tracking system is capable of tracking position data up to four trackers, this is not currently implemented, and will be pursued in a future investigation. Use of directional audio cues have been explored very briefly, but needs a detailed study.

References

AIR230 Course. Royal Naval School of Flight Deck Operations. See also Windows User Guide.

Hand, C. (1997). A Survey of 3D Interaction Techniques. *Computer Graphics Forum, 16 (5),* 269–281.

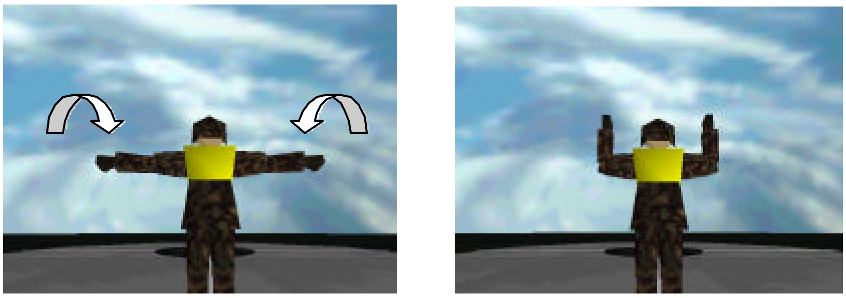
Kalawsky, R.S. (1993). From visually coupled systems to virtual reality. An aerospace perspective. *Proceedings Computer Graphics, 91*.

Nemire, K. (1998). Individual combatant simulator for tactics training and mission rehearsal. *SPIE Proceedings, The Engineering Reality of Virtual Reality, 3295,* 435–441.

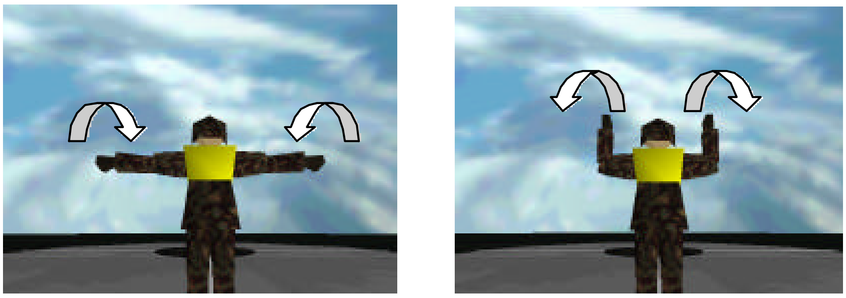
Mapes, D. & Moshell, J.A. (1995). A Two-Handed Interface for Object Manipulation in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 4 (4),* 403–416.

Mine, M. (1997). *Exploiting Proprioception in Virtual Environment Interaction.* PhD Thesis, University of North Carolina, Chapel Hill, Computer Science Technical Report, TR97-014.

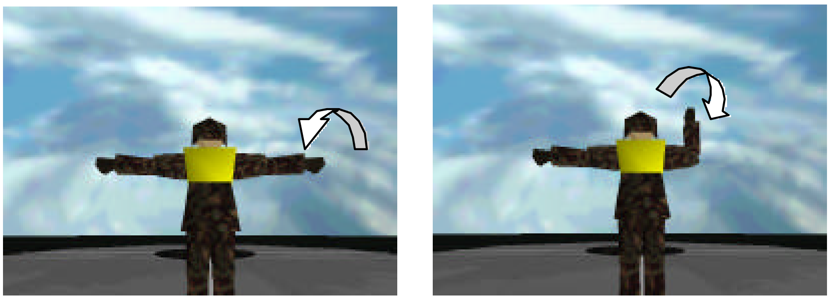
Trott, E.A. (1999). *An Investigation into the use of Virtual Reality for Naval Flight Deck Operations Training.* MSc Project report, Cranfield University, Royal Military College of Science, Shrivenham.



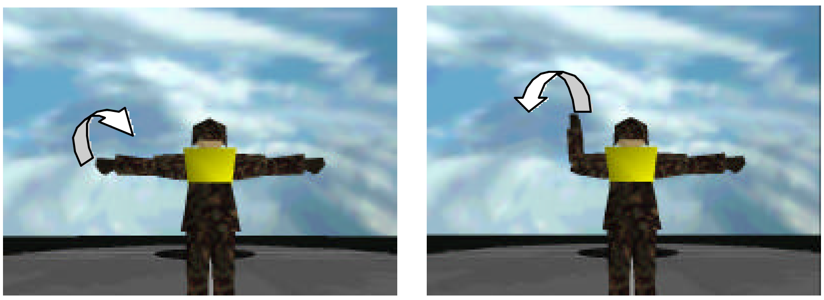
FDO Ready to receive – you are cleared to land



FDO Approach

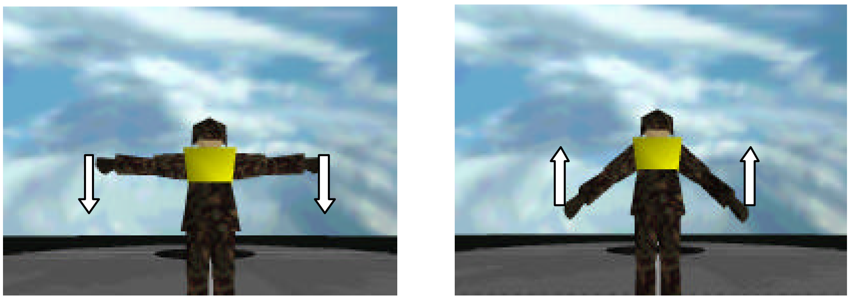


FDO Move left

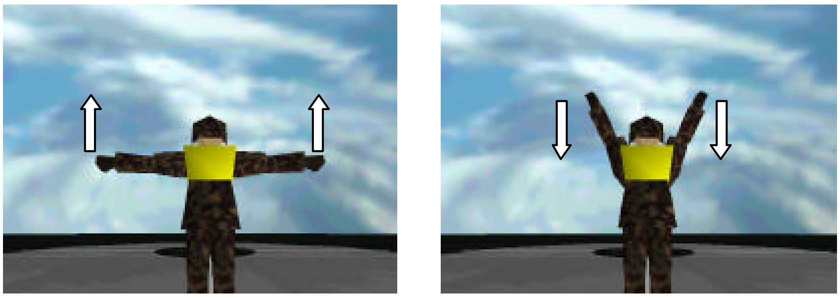


FDO Move right

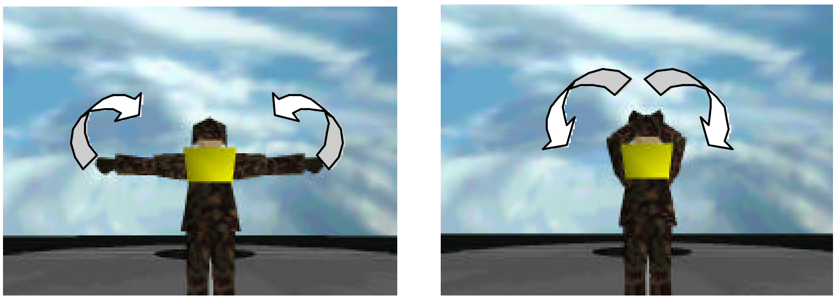
Figure 3: A subset of Flight Deck Officer’s hand signals



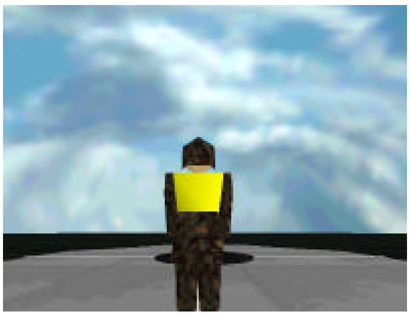
FDO Down



FDO Up



FDO Wave-Off



FDO Hold

Figure 4: A subset of Flight Deck Officer's arm signals (ctd.)

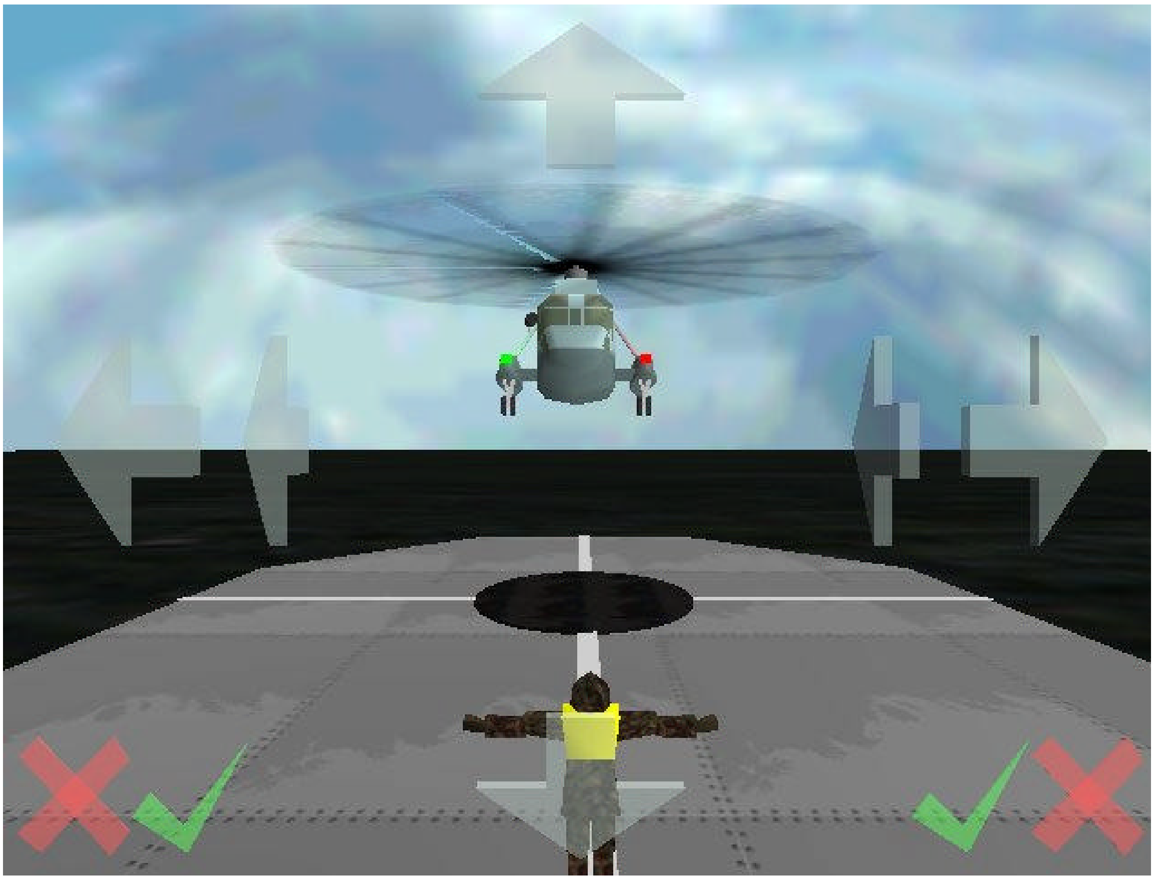


Figure 5: Virtual menues used in the virtual environment

Mission Debriefing System

Major Birger I. Johansen[1] and Bo Fredborg MSc EE
Senior Systems Consultant and Systems Engineer
Systematic Software Engineering A/S
Søren Frichs Vej 38K
8230 Åbyhøj
Denmark

Abstract

Systematic has developed a debriefing system for aircraft crews to improve their skills based on experiences from completed missions. The system is developed on Commercial Off The Shelf (COTS) software and on a PC. The panel should see this input as a portable, low-cost Virtual Reality (VR) training system for aircraft crews. The benefit of the portability is that the system can be **used anywhere** the unit is deployed and **by any crewmember.**

Flight hours are rather expensive and therefore the air forces must maximise the benefits from spent flight hours. This, combined with the fact that most air force units need to operate from different deployments remote from home bases, led the operational fighter squadrons to express a need for a low-cost debriefing system.

The users were directly involved in the design and the focus was set on functionality — not technology. This approach has resulted in a system which gains accept among users and therefore becomes an everyday training tool. Driven by user requirements, the system is developed to run on a Microsoft Windows 2000 platform, and the system can interface with other systems. Furthermore, it has been essential to develop a system, which could be rapidly implemented.

The debriefing system uses already existing information from the aircraft. The aircraft is equipped with Global Positioning System (GPS), three video cameras, and a microphone system to record the pilot's voice communication. The video cameras record the pilot's view through his head-up display and the entire instrument panel.

Prior to the debriefing session all information from the aircraft (GPS-data, video- and audio recordings) is fed into the debriefing system. The GPS-data is loaded into a three dimensional (3D) model containing geographical information, the video and audio recordings are digitised, and all data are synchronised. On each monitor, four visual sources can be displayed concurrently, e.g. video recordings from three different aircraft and the graphical 3D view of the area, including aircraft. The selected visual sources are displayed along with a selected audio recording. The 3D graphic makes it possible to see and follow selected aircraft from different perspectives on their mission. Furthermore, it is possible to see them chase other aircraft and to track their route by position, direction, and speed. The crew and other mission participants can by themselves prepare and execute the debriefing session.

Systematic has developed a portable, low-cost VR training system for aircraft crews, which converts reality to virtual reality, reflecting the reality. The chosen approach, with heavy user involvement, has resulted in a system, which is easy to use and will gain much better acceptance. A system based on well-proven COTS products reduces costs as well as risks. Finally, the system gives added value to the flight hours spent.

Introduction

It is our aim with this paper to disseminate understanding for the possibilities given by new commercial off the shelf (COTS) products — in this case especially for low cost virtual reality training tools. We find that today's COTS software fulfils most of the requirements that the military has to an everyday debriefing system. By combining the COTS products using Systematic's competence in software integration, a low-cost easy-to-operate operational training system, has been developed.

Through this paper, we discuss functional requirements, use of commercial state of the art technology, influence on training and human performance requirements, and describe the development process and functionality in our debriefing system.

In connection with the training of combat pilots much time is spent on manoeuvres in actual air combat techniques. The Danish Air Force spends more than half of the flight hours on such manoeuvres. Furthermore, the remaining flight hours often contains elements of air combat. It is therefore essential to get full benefit from the training, especially as flight hours are extremely costly. Nevertheless, subsequent debriefing and evaluation of a training session is often deficient or non-existing.

The present project has endeavoured to remedy this inadequacy by investigating the possibilities for building an inexpensive, simple and user-friendly, but yet high-tech, mission debriefing system, for "everyday use". We have used virtual reality (VR) and 3D techniques for

[1] For correspondence with author: bij@systematic.dk, tel. +45 89 43 21 53; with company: tel. +45 89 43 20 00, fax +45 89 43 20 20

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

constructing factual conditions for training in a Virtual Environment (VE). The VE facilitates the debriefing of pilots and thereby enhances the learning. Presently the system is developed as a 1. generation version with basic functionality financed by our company. We find though, that the idea has much potential, and we will promote our ideas broadly within NATO.

Background

Security in the Euro-Atlantic area has substantially improved during the 1990s, by comparison with the four decades that preceded them. The threat of massive military confrontation has gone, and co-operative approaches to security have replaced former confrontation. Nevertheless potential risks to security from instability or tension still exist.

In these changed circumstances affecting Europe's security, NATO forces have been adapted to the new strategic environment and have become smaller and more flexible. Conventional forces have been substantially reduced and in most cases so has their level of readiness. They have also been made more mobile, to enable them to react to a wider range of contingencies; and they have been reorganised to ensure that they have the flexibility to contribute to crisis management and to enable them to be built up, if necessary, for the purposes of defence. Increased emphasis has been given to the role of multinational forces within NATO's integrated military structure. Many such measures have been implemented. Others are being introduced as the process of adaptation continues.

Airforces are characterised by their ability to operate from far distance, geographically dispersed bases and concentrate their efforts against the main targets. They are also able to react very fast and to maintain a high degree of readiness. These characteristics have made airforces even more important to NATO's new strategic concept, Combined Joint Task Forces (CJTF). The main issue in this concept regarding air forces is high readiness, interoperability and the ability to operate away from home bases with a minimum of preparations. Furthermore each participating unit must be able to perform a larger variety of roles, than before — e.g. using heavy bombers for close air support. The operational environment has become much more dynamic — it is never possible to foresee which type of operation that will turn up. This again puts higher demands on continuous and flexible training.

Another consequence of the new operational environment is the reduced military budgets. This means that it is essential to gain as much as possible from the applied training efforts. In real operations like Allied Force in Kosovo last year, it is extremely important that the pilots learn from each mission to make continuos improvements. In this specific example, most of the participating units operated far away from home bases. Therefore they were not able to take advantage of their usual static training equipment and simulators.

Project Objectives and Means

An obvious need for mobile training, rehearsal and debriefing systems has evolved. Given the fast development within virtual reality technology and low cost flight simulators for PCs, we have seen a good opportunity to use commercial technology and existing sensors, video recordings, and tapes from the aircraft to develop a debriefing system for air force pilots.

The overall objective was to create a low cost, easy-to-operate, and transportable debriefing system. With this objective the intention was that each training session and live mission should be followed up by a high-quality debriefing activity, giving full benefit of the costly flying time to the pilots.

The aim was to base the system on COTS products and existing and electronically available data from the aircraft. Furthermore, the aim was to combine the collected data from the aircraft and thereby constitute a 3D Virtual Reality (VR) replay of the completed missions and training sessions.

The project is financed by Systematic and the ingredients used are Systematic's skills and knowledge, technologically as well as military, a range of COTS products, and the requirements set by airforce pilots.

System requirements and functionality

This section describes the scenarios and missions that are supported by the debriefing system. To stress out the need for a debriefing system, we give a brief description of the main categories of existing systems.

Air Combat Manoeuvring (ACM)

Air combat comprises all kinds of manoeuvres in the air in a one-to-one, many-to-one or many-to-many situation. ACM includes all the classical movement patterns such as half loop, full loop, split S, break turn etc. Basically, air combat is a question of gaining the right position in relation to the opponent.

A training session consists of a number of scenarios, ranging from 3 to 10 — depending on the number of fighters involved. A scenario lasts from 5 to 10 minutes. The starting point of a scenario is an initial position where, for example, the different players have got radar contact (approx. 30NM distance). Typically, the situation then develops rapidly, depending on the actions that take place during the session. After only a few minutes, the situation typically becomes very complex and the pilots often lose control of the situation. As an example, a pilot who tries to escape will lose control of what is going on, as he has no longer radar contact with the other fighters.

When a training session is over, the pilots involved should evaluate the session. This is typically a difficult process, partly because the individual sessions develop in a complex way where each pilot may have different opinions on what actually happened, if they are able to contribute to the situation at all. But also because the individual scenarios become indistinguishable when the pilots have returned to the air base. As a result hereof, debriefing is deficient or non-existing. Consequently, much value of the training is lost. This should be viewed against the large resources spent on keeping the fighters in the air.

Existing Solutions (ACM)

In order to enhance the debriefing possibilities, various systems are available for the pilots for recreating the individual scenarios that constituted the training. Generally speaking two solutions exists: A low-cost and an expensive solution.

Low-Cost Solution: Video

The F-16 fighters used by the Danish Air Force are equipped with three standard video cameras, which records the Head Up Display (HUD) and the two Multi Function Displays (MFD). The pilots can use these videos in a subsequent debriefing. Videos are excellent for the initial scenario and evaluation of shootings. In a debriefing situation, the pilots involved will endeavour to recreate the individual scenarios in the training session. If the pilot has lost control, however, videos are of little use (the radar image may be of no value). Furthermore, it is difficult and time-consuming to synchronise multiple videotapes and ECM as well as kill removal are not covered by video at all.

The Expensive Solution: Real-time ACM Instrumentation (ACMI)

Real-time ACMI covers the expensive and extensive solution where the individual fighters that participate in the session downlink information in real-time to a control station on the ground. Via the control station, the individual scenarios are monitored and stored for later debriefing. The control station may even intervene during the training session, either in order to influence the situation in a certain direction or due to kill removal.

Real-time ACMI involves pod-mounted electronics (GPS, MUX-BUS interface and data link) as well as antenna coverage on the ground and all control facilities on the ground. Consequently, the solution is quite costly in terms of electronic equipment and staffing, and ACMI will not become a natural part of every training session. ACMI must be planned a long time in advance and will only be used few times a year.

Systematic's mission debriefing system

Based on informal discussions with both pilots from Air Station Ålborg and the Danish Air Materiel Command, we have developed a first generation model to show the possibilities.

The first generation of the debriefing system is an autonomous system and does not require any changes in the cockpit or instrumentation of the aircraft. The system is centred on a debriefing facility, based to the greatest possible extent on COTS hardware and software.

The debriefing system uses already existing information from the aircraft, the Global Positioning System (GPS) data, the three videos (HUD and 2xMFD), and a recording of pilots' voice communication.

The HUD, MFD, voice recording, and GPS data of the individual aircraft are loaded into a Personal Computer (PC), which synchronises the data. From the synchronised data the PC constructs a 2D/3D synthetic world of "what happened".

The three videotapes and the voice recording are used to give a detailed image of the pilots' actions, displaying what happened inside the cockpits. The GPS data from all aircraft are loaded into a 3D model of the battle cube. The 3D model does, just like a Geographical Information System (GIS), contain a 3D graphical model of the landscape in the battle cube. This 3D model of the landscape combined with the aircraft GPS data gives a "Gods eye view" of the battle cube. The debriefing system makes it possible to navigate around in the battle cube. This makes it possible to view the scenery from different perspectives.

All aircraft that can provide the information described above can be included in the debriefing session. Consequently, the system can be used not only by the Royal Danish Air Force's F-16 fighter pilots. Furthermore, a debriefing system like this can be used independently of the geographical location and extension of the individual training sessions. Compared with the real-time ACMI system, this provides an obvious advantage; the real-time ACMI system is not mobile, but limited to the location that is covered by the antenna equipment of the ground station.

The latest techniques in Virtual Reality and 3D have been investigated in connection with the construction of the synthetic world. These areas undergo extensive research and development within the experimenting field of computer science, and are consequently considered to contain some of the building blocks for the future development within HCI (Human Computer Interaction). The debriefing system includes leading edge technologies within these fields. It is our aim to present a system that will delight and motivate the pilots to carry out high-quality debriefing.

Development of the debriefing system

This section is a brief description of our approach to the project. Based on our interviews with potential users, a retrieval of user requirements and a study of existing COTS products and their facilities, we started the development process. Knowing that we had to do with

new technology, it was essential for us to study and develop small prototypes of the different functionalities in the system. We decided to break the system into three main subsystems, which were to be developed and tested sequentially. The initial aims were:

- To see if we could develop 3D graphics using “cheap” COTS technology and already available data.
- To test the different 3D graphical components/effects that we wanted to make use of.
- To establish a 3D-terrain model, which was suitable for debriefing purposes.
- To establish a user-friendly interface and the framework from which the debriefing application should be prepared and presented.

In the following text we describe each of the initial prototypes, it’s purpose, the method used to develop it and the result/experiences gained.

Prototype 1

This part resulted in a 3D-terrain model with a visualisation of a number of aircraft including their tracks, so that one can get an overall view of the full mission or extracted parts of a mission or flight.

Purpose

- To get a 3D-terrain model and to show it on a PC (We decided to get the necessary data from the Danish F-16 simulator).
- To make a 3D visualisation of aircraft (including their historical tracks).
- To create lively navigation and animation methods (the aircraft should be able to manoeuvre and navigate in a realistic way so that an aircraft would bank naturally when turning and so forth).
- To enable the user to choose between different angles of view (e.g. “God’s-eye-view”).
- To visualise other objects (e.g. Surface to Air Missile sites with threat domes).
- Portability: To be able to port the system between the normal PC platform and a more static SGI graphical supercomputer with holobench.

Method

In brief we have had a very open and innovative approach where following main activities were carried out:

- Information search on the Internet to get components and pieces of code, which could be useful.
- Selection of a portable visualisation core component. (Optimizer™ from Silicon Graphics).
- Courses in the use of Optimizer™.
- Prototyping and test using visualisation methods and navigation.
- Get inspiration through the studies of existing ACMI systems.
- Initial development on PC — later ported to SGI.

Results

These were the results we got from our first developments:

- Functionality to convert the database from the F-16 MLU simulator to “PC-format”.
- A prototype application showing a landscape of size 10 x 10 NM.
- Playback of flights. (Specifically two flights flying different routes.)
- Possibility to see the flights in a follow-mode (seen from one of the flights or in a “God’s-eye-view”).
- Portability between PC and SGI (holobench).
- Possibility to run the application on a PC with a powerful graphics card.

Prototype 2

The next step was to develop an application that could visualise a complete geographical database and to make 3D movement through the landscape.

Purpose

- To create functionality to visualise a complete geographical database covering a normal theatre of operations.

Method

- Use experiences from prototype 1.
- To develop and implement efficient methods to get and drop tiles of terrain in the visible area.
- To convert the F-16 MLU simulator database to PC-format”.

Results

- A prototype 2 application with functionality, which in principle (if terrain data is available) can show any given terrain.
- Geographical data enabling the system to cover Denmark and Southern Norway.
- This prototype was only developed for a PC.

Prototype 3

The third prototype is the set-up and administration tool, developed on a Microsoft Outlook user interface.

Purpose

- To obtain functionality to administrate flights and missions. (A flight is an operation/flying session performed by one aircraft and a mission is a combination of concurrent flights).
- To be able to perform video playback.
- To synchronise video inputs and the 3D-terrain model.
- To present a graphical user interface (GUI) for debriefing and administration in a Microsoft Outlook view.

Method

- Standard components were to be used
  - Standard Template Library (STL) from Silicon Graphics.
  - Microsoft Foundation Classes (MFC) from Microsoft.
  - Windows media standard components for video playback.
  - Microsoft Access Database.
- Use of simple application development (Visual C++, 6.0).
- Use of well-known components for the GUI (Microsoft Outlook).

Results

- A quad-view (four concurrent views on same screen) with an intuitive timeline that permits playback, review, pause etc.
- An intuitive, easy-to-learn GUI.
- Use of Windows standard functionality to synchronise video and data.

Integration to a first generation model

Before integration of the three prototypes into the first generation of debriefing system, we had to solve some minor problems that occurred during test of prototypes:

- Geographical data are extensive and requires a harddisk of at least 1GB for the database. We improved our hardware to the necessary level.
- Movements through the 3D terrain require loading and initialising of huge amounts of data. Therefore a dual processor system and a very fast harddisk must be used to give video and other resources enough processing power.
- Using video and 3D-graphics in the same session creates performance problems — Windows 2000 combined with multiple graphic cards solves the problem.

Once these problems were solved we were able to load the real, digitised video from F-16 aircraft and through prototype 3 we could initiate, administrate and run the debriefing system with the introduced functionality. The result is promising and after some pre-tests with real users and the necessary adjustments and improvements in functionality a flexible system is ready to be implemented with operational fighter squadrons.

Use of Systematic's debriefing system

Using the Systematic debriefing system is a 3-step process:

- Digitisation of source data
- Mission/flight set-up
- Debriefing

These processes are described in the following.

Digitisation of source data

After completing a flight, data collected from the plane must be converted to formats suitable for computer processing. Analogue data must be digitised and stored in appropriate formats.

- **Video.** Video recordings from the HUD and MED's must be digitised and converted to "mpg1" format.
- **Discrete flight path information.** Flight path information consisting of at least position (time, latitude, longitude, height) and optionally orientation (heading, pitch, yaw).
- **Event registrations.** Identification of events that occurred during the flight. These could be: Weapon-release, radar lock-on etc.
- **Environment.** Stationary and moving objects which give important input to the flight debriefing. This could for example be location of a SAM site.

Mission/flight Set-up

The purpose of the mission/flight set-up phase is to arrange the source data into logical units such as flights and missions. For example, a *flight* is a container for all data relating to a flight including; name of the pilot, identification of the plane, the videos recorded from the plane and the flight path data from the plane.

Data is arranged in a hierarchical structure:

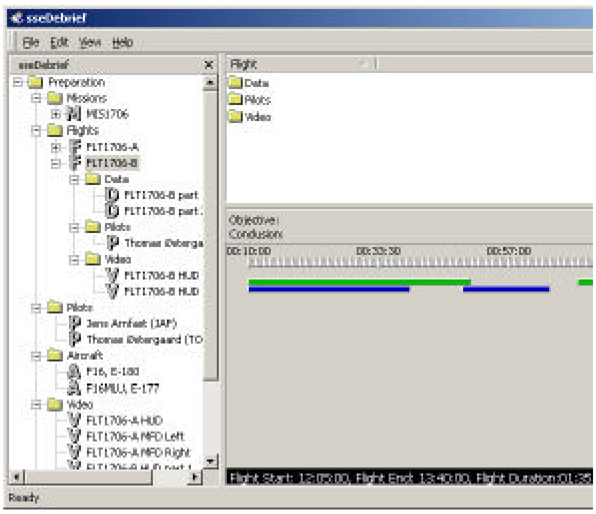


Figure 1: Data structure

The different types of data/files should be read as follows:

- **Mission.** A mission defines a collection of related flights. A debriefing typically involves several flights.
- **Flight.** A flight defines the pilot, the plane, a set videos recorded from the plane and flight path recording from the plane.
- **Pilot.** Defines the characteristics of a pilot
- **Plane.** Defines the characteristics of a plane/aircraft. Aircraft type/model, visual representation

- **Video.** Defines video recording from a plane (related to a plane). Includes start and stop time for the video.
- **Data.** Defines flight paths.

Debriefing
A debriefing is concentrated around a mission.

The screen is divided into five sections as displayed in Figure 2. Four of the sections are dedicated to displaying video and/or the 3D synthetic environment. The remaining section is dedicated to the timeline and playback controls.

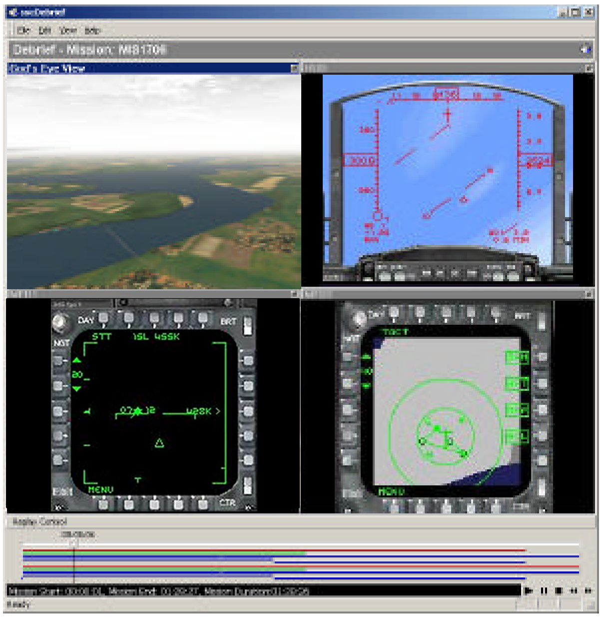


Figure 2: Division of screen in debriefing mode

The user can make use of following functionality:

- **3D syntetic environment:**
  - God's-eye view
  - Follow mode
  - Free movement
- **Video-playback:**
  - On/off
  - Sound
- **Play-back control:**
  - Play
  - Fast forward
  - Reverse
  - Slow-motion
  - Single step
  - Search (time, event)
- **Pop-up time based annotations/Attachments on:**
  - Data
  - Audio/Video
  - Flight
  - Mission
- **Computed "annotations":**
  - Speed, Height
  - Distance
  - Radar coverage
- **Information layers (on/off toggles)**
  - Flights

Combining commercial off the shelf (COTS) technology with military requirements

To reduce cost and improve the usability and learning process, Through studies of a range of commercially available products, we have experienced that today's COTS products basically cover all given requirements to a debriefing system.

COTS Hardware
The PC market, driven by the requirements set by the entertainment industries "need" to produce more and more realistic games, produces high-performance affordable systems. Current state-of-the-art entertainment PCs are capable of delivering the high performance in the areas essential to 3D graphics and video applications. The essential areas are:

- Processing power — Fast processors are required to handle movements through the 3D-terrain model. Multiple processors are recommended.
- Main storage — Memory (RAM) is essential to store the 3D-terrain in use.
- Mass storage — Harddisk space is needed to store digitised videos and 3D synthetic terrain. Today mainstream harddisks are both fast and has large capacities.
- 3D graphics — A 3D accelerated graphics adapter is essential to produce 3D synthetic environments at suitable resolution and frame rates. The entertainment industry drives the need for 3D graphics performance. Current and next generation consumer 3D graphics systems are powerful enough to drive the 3D synthetic environment.

COTS Software
We have found that most of the necessary software for the debriefing system is available in different COTS products, which can be acquired within a reasonable price or directly downloaded from the Internet. By using these products we also make it easier for the user to learn to use the system. We decided early in the project to use Windows 2000 instead of Windows NT. The reason for this is that Windows 2000 can handle concurrent use of video and 3D-graphics.

Experiences
We have spent many hours searching for relevant software products on the Internet and other places. We have certainly gained benefit from these efforts. Generally speaking there is COTS technology available — especially from the entertainment industry — to support and develop a range of high-tech, virtual reality training systems. Our task has almost been reduced to integration of already well-proven and tested blocks of

software code. However, it must be stressed out that the main challenge was to make the individual products work together.

Perspective

The debriefing system has great extension possibilities. As an example, the air force's simulators could use the debriefing system for evaluation of the simulation training. By doing so, simulation and use of the debriefing system will become an integrated part of the general simulator training. Consequently, the possibility for evaluating "what if" situations (situations where a training scenario is evaluated against *new* actions) would become a reality. An existing training scenario that has been practised and debriefed in the debriefing system could provide input to the simulator. The simulator could then fly with the scenario, and what-if situations could be simulated in order to evaluate the effect.

The interaction with other ground systems, such and C3 and Mission Planning Systems, are further areas to look into. As an example, the debriefing system could be used to build an Airspace Co-ordination Order (ACO): With a "magic wand" the operator could guide and virtually draw a route through the 3D landscape. An F-16 fighter could then use the ACO generated. When the mission is completed, the route planned and carried out could be compared in the debriefing system.

Another opportunity would be to investigate the debriefing facility in an interaction with other armed forces. As an example, the Navy's air combat system could be investigated and evaluated. One of the problems of the Navy in air combat is finding the optimal defence process, and the debriefing system may turn out to be useful.

The F-16 is equipped with a MUX-BUS interface. Through this interface much more information, e.g. weapon-release can be accessed. Recording these information and successively replay during the debriefing will give a much more detailed image of the flight.

The opportunities described above are just some areas where it may be possible to use the debriefing system. When the system is in operation, other opportunities are likely to appear, and technology will show us which.

Conclusions

We have developed a mission debriefing system that in principle covers the basic requirement and to some extend even exceeds these requirements. No dedicated software has been developed for use in this first generation of the system. The input to the debriefing system is not made especially for this purpose, but already available sources have been sufficient (a digitisation of the flight videos has though been necessary). Available COTS software and hardware has shown its value for this purpose, which means that the main task for us has been to integrate already available products and input. As integration is one of our company's main business areas, we are able to do this quite fast and therefore within an affordable price.

This page has been deliberately left blank

Page intentionnellement blanche

Mine Clearance in a Virtual Environment

Laurent Todeschini, Thérèse Pasquier, Pascal Hue and Paul Gorzerino
DGA[1]
Etablissement Technique d'Angers
route de Laval, BP 36
49460 Montreuil Juigné
France

Abstract

At the same moment as France completed destruction of its stock of anti-personnel mines (21/12/99) in accordance with the 1998 Ottawa Agreement, in more than 60 countries there were 100 million live, buried "permanent sentinel" mines continuing to mutilate the inhabitants of mine-infested regions, most of the wounded being children (600,000 people affected over 20 years, one person killed every 20 minutes by these devices designed to terrorise civil populations during the war, whose effects persist for a long time afterwards). Paradoxically, confronted by the sophisticated manufacturing techniques of these "cowardly weapons", French sappers use a rudimentary mine clearance technique to render zones viable for the civil population. With the aid of a bayonet-type tool, the operator probes the ground until he hits a suspect device. This task is carried out blind and one of the problems is identifying the presence of a mine and distinguishing it from a false alarm. This technique, demanding 100% results, based on the skill and experience of the mine disposal team, is taught by the Minex Centre of the Applied Engineering Applications College.

The Human Factors division of ETAS (Etablissement Technique d'Angers), a part of the DGA, has built and tested version 1 of a demonstrator and virtual environment for teaching this technique. One group under training now has been able to distinguish the methods for discriminating shapes after several contacts of the probe with the mine.

In its version 2 (addition of force feedback), this demonstrator has become a genuine teaching tool for mine clearance strategy, enabling the instructor to validate the relevance of the students' probing, to minimise the amount of probing and therefore to increase the reliability of the decisions during an actual operation. In due course, this tool will also enable the technique to be taught to civilian populations and thus accelerate the process of decontamination which still takes a long time, costs a lot of money and, especially, costs lives.

Technology development is already enabling us to consider version 3, a portable system which uses mathematical analysis of the probing geometry during real operations, and by comparison with a database, offers genuinely enhanced assistance to making decisions and taking action.

1. Problems of Mine Clearance

The difficulty of mine clearance is that of DRI (detection, recognition, identification) associated with some action.

The main problem is detecting the device: the mine clearance expert probes the ground in a systematic manner in a 5×3 triangular grid arrangement to try and *detect* the presence of a foreign body. If the probe hits something, the mine clearance expert halts his movement. He then probes in order to discover the extent of the object and to determine its shape, which will enable him to *recognise* the presence of an object. He then clears away the soil covering the object and *identifies* it as being a mine or not, a munition or some unknown device.

If a device is present, a specialist intervenes who, after having made a detailed identification, detects any booby traps, analyses the condition and mode of triggering, and decides to deal with the device either by destruction or by rendering safe.

This paper is only concerned with the DRI task in the initial phase; it is a difficult task, performed blind, necessitating the mobilisation of sensors in seeking stimuli which are indicators both for the accomplishment of the task and for a perfectly controlled motor activity; indeed the relevance of the indicators is dependent on the steadiness of the prodding (the angle of incidence of the probe must remain constant), this angle is a safety factor and makes it possible to attack the mine at its edges and not from above where the initiator is generally situated. (angle of incidence lies between 30 and 35°).

2. Aim of the Research

In order to design a simulator for teaching manual probing specific to the activities involved in mine clearance, the Human Factors division of ETAS, in association with the Angers Cognitive Psychology Research Laboratory, engaged in research into learning conditions in a virtual environment. This research was the subject of a thesis entitled: *Influence of sensorial methods and individual characteristics on the conduct of target detection in compared environments: the case of virtual and actual environments.*

The conduct of sensorial learning was observed for a task in a real environment and then in a virtual environment. The experiments were conducted in real

[1] For information, phone +33 (0) 241 936 831.

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

time and took three aspects into account: the detection of shapes; the detection of textures; and the detection of sounds. Detection was based on visual, kinaesthetic and auditory indicators. The experiments conducted in a virtual environment were limited to a shape-detection task.

Nowadays, virtual reality based on visual immersion makes it possible to explore and augment visual information in an enhanced manner and it also allows, at a lower level, transmission of multi-modal sensorial information by haptic, tactile or auditory feedback systems, close to those felt in real life.

For qualitative and financial reasons, this thesis was restricted to a part of the significant information of a blind, sensori-motor task. The mine clearance expert's task was not reproduced identically. A selection was made of the sensorial modalities of exploration. Emphasis was placed on the guiding of movements aiming at the target using visual and auditory indicators excluding any use of a haptic interface. Therefore the strategy of probe movement was tested in virtual space:

1. by observing and analysing a shape detection task;
2. by comparing the learning of this task when conducted in a real and in a virtual environment.

3. Development of Virtual Reality and Research in a Multi-disciplinary Team

The number of articles concerning new virtual reality interfaces shows the diversity of applications and their development from the leisure domain into the professional world and firmly establishes virtual reality as man’s new environmental tool.

What effect does confrontation with the virtual world have on human behaviour? Is such immersion neutral, or does it influence behaviour or generate different behaviour?

Scientific research was directed towards specific prototypes for learning in a virtual environment with the aim of preparing the user for a new real environment and to encourage his adaptation by devising a new generation of training facilities.

Few multi-disciplinary scientific teams incorporate experimental psychologists, cognitivists and neurophysiologists for studying the perceptive, sensorial and cognitive effects on man in the virtual environment and reflect, *inter alia*, on man’s ability to transfer learning from the virtual world to the real world. Such teams are still not very numerous today (Kalawsky, 1999 and Fusch, 1999) in spite of a strongly expressed need to integrate human factor dimensions, both cognitive and sensorial, into technological research.

The development of such collaboration is becoming urgent in the field of man–machine interactions in order to accelerate the development of knowledge and understanding of the virtual reality interfaces which are still limited today. Evaluating the “human factor” component is complex and covers numerous aspects such as performance linked to sensorial capacity and cognitive processes.

It was from this viewpoint, multi-disciplinary team and performance evaluation, that the Human Factors division of ETAS became interested in virtual reality and developed a virtual reality platform with the aim of testing the upper level interfaces in learning specific movements.

4. Experimental Conditions

4.1 Selection of subjects and the experimental task

Experiments were conducted with a group of 27 subjects, all of whom were mine clearance experts. The task was that of target DRI. This simple gestural movement was defined as an activity of blind probing aimed at identifying the structure of a hidden target, using a probe, an intermediate tool extending the operator’s hand. The mine clearance expert had to identify three shapes of concealed targets which were rectangular, triangular and round. The mine clearance expert probed into a container placed in front of him, respecting the angle of probing used in mine clearance (between 35 and 45°) and had to identify each shape several times in succession in order to measure the effect of learning.

The mine clearance experts were split into two subgroups one of which had the benefit of a visual aid (a special virtual reality feature which made it possible to show impacts on the target) while the other did not. The chosen variables were the detection time, the number of probings and the quality of the response.

Virtual reality made it possible to display the incidence of the probe with respect to the terrain and to monitor constancy.

4.2 The conditions for exploring the virtual environment

The subject wore a helmet with stereoscopic vision, a V6 from Virtual Research, each channel being connected to a graphic map so that the image was retransmitted in 3D. The mine clearance expert’s real probe was used to reproduce direct contact between the hand and the exploration tool. The position of the helmet and the displacement of the probe were controlled (6D) by “flock of birds” position sensors.

The frame of reference was established as a function of the environmental context, which undergoes significant changes in the virtual environment, and the subject's visual capabilities. In order to construct the virtual environment special attention was paid to the selection of the essential markers for forming the frame of reference:

a. at the visual level, according to the concept of identifying the shape and to remain faithful to the analysis of a simple task, the visual markers had basic geometric shapes concealed in an environment with filtered geometrical data and colours. Visual representation of the size of the probe was proportional to its actual size.
b. at the haptic level, according to the concept of transfer of sensorial modalities, the markers were partially transformed into auditory markers. The collision detection points were signalled by sounds which symbolised the times of contact between the probe, the environment and the targets. For gestural

guidance and accuracy reasons, the real medium was replaced by a substitute real homogeneous medium in which the mine clearance expert carried out the probing.

c. at the kinaesthetic level, the constituted virtual environment gave the opportunity of traversing the walls and therefore influenced the kinaesthetic and proprioceptive movement of the subject, who lost the notion of rigidity of the wall.

5. Theoretical Approach

5.1 Sensori-motor and cognitive domain

The observations of this study of a shape detection task were conducted in a sensori-motor learning context. In the real environment, to read spatial information the sensori-motor act is associated with the subject's cognitive capabilities (Paillard, 1985). The treatment of spatial information cross-refers to the detection capabilities and therefore to the attention the subject applies to discriminating sensorial space. This space is a function of the information reflected by the environment. Thus, the sensori-motor act is linked to the attention capabilities of the subject and to the mental loading due to processing the information received from the environment. The subject's performance will also depend on the mental representation of the action undertaken. This information processing occurs in three stages:

- a perceptive stage which corresponds to processing the stimulus;
- a motor stage which is the transmission of the action undertaken on the medium;
- a response processing stage which is the subject’s stimulus–response translation.

The identification of an object is the subject of a multi-modal processing (visual, auditory, kinaesthetic and tactile). Similarly, a subject may recode information under several sensorial modalities. However, in order to identify an object each individual will recode according to his particular sensorial predisposition (Ohlmann, 91), which enables the different approaches to be differentiated. Research has emphasised the interactions between the various modalities and the major implications in co-ordinating sensorial activity. Study of the relationship between perceptive systems describes a variation of the predisposition of perceptive systems according to the object of the study (Hatwell, 1994).

The individual mobilises “decoders” as a function of the data to be extracted and of his sensorial capabilities in processing information. Recognition of the shape of an object or stimulus is defined by the object’s specific intrinsic characteristics (by its shape, dimensions and colour) and by its extrinsic characteristics (its position and orientation in space).

Given that the visual dimension is predominant in the virtual environment and given that the priority of perceptive systems can change according to the type of task in the real environment, can this perceptive priority be modified in passing from one environment to the other?

How does the individual process information when immersed in a virtual environment? Are the cognitive processes employed in the real world automatically efficient when the person is immersed in virtual reality? The work of Morineau, Boujon, Papin and Le Bouedec (1996) tends to show that the adult plunged into a virtual world for the first time appears to use the cognitive processes coming under the preoperative structures of a 5-year-old infant. These results project the idea that immersion in the virtual world requires acclimatisation or learning.

The cognitive dimensions of the personality may intervene in processing information and have been the subject of numerous papers (Huteau, Marendaz & Ohlmann). In this context the concept of “dependence and independence with regard to the visual field” offers relevant explanations in the real environment.

The DIC (dependence and independence with regard to the visual field) is a theory on the personality factors presented among cognitive styles referring to the work of Witkin (1948). Exploration strategies differ with the IC (independent with regard to the visual field) and DC (dependent with regard to the visual field). Huteau (1985) developed the theory of the DIC and qualifies the IC by higher discriminative capacities and level of vigilance, basing this on egocentric factors and their own perception built on gravitational, proprioceptive or kinaesthetic factors, whereas the DC use more visual factors for referencing themselves in space. They will be very attentive to the positions of others, referring to external factors. Ohlmann and Marendaz (1991) studied this same theory from perceptive conflicts.

Before action, the operator employs a conduct, a manner of proceeding and of giving a reasoning whose degree of complexity varies with the task. In addition to environmental factors, personal factors and the specific nature of the action are going to influence the subject.

The reference point of this study relates to the concept of restricted spaces in a static situation. The subject relies on his capabilities of spatial representation. In order to recognise a shape and characterise it the subject must be capable of selecting stimuli that can be arranged in a simple or complex fashion.

These will be differentiated by combinations of specific information about a shape whose basic identifying markers will be based on arrangements of points characterised by the distance separating them, their orientation, intersection and movement. The subject will mobilise his attention to create grouping factors so as to determine the boundaries and define the contours. In order to perceive a shape and to construct a representation, each subject has need of information which may be total or partial. The person's strategy is based on a representation using part or all of the constituents of the shape.

Image processing cross-refers to perceptive models of basic features to guide a discriminatory behaviour

between global information and more analytical local information. In differential psychology, cognitive styles are evoked by global or analytic strategies in analysing various activities such as learning, memory attentiveness or games strategies.

5.2 A specific feature of the task: remote manipulation

As described above, the mine clearance task is performed blind. The target is masked off from the visual field and probing is carried out with a tool, the probe. The probe is a link enabling three types of sensorial information to be transmitted:

- visual, by the presence of marks left on the surface of the soil enabling the shape to be identified;
- tactile, which makes it possible to detect collisions and identify textures;
- auditory, during collisions for identifying materials.

The problem is to correlate these various sensations.

The mechanoreceptors situated in the hand and at the ends of the fingers possess perceptive acuity which is strongly discriminatory and makes it possible to decode the detailed information which is characteristic of the objects dealt with. Is the acquisition of information as powerful when the hand is not in direct contact with the object?

Recent studies on professional situations where interaction with the world necessitates an intermediary contact object were aimed at measuring the performance of haptic spatial recognition in a real environment (Lederman & Klatzky, 1998). This work was aimed at providing information on tactile manipulation of intermediate interfaces in remote control or in virtual environments.

5.3 The rapid development of the virtual environment

The rapid development of technological and computer facilities over the past few years has resulted in a possible skewing between the initial analyses and recent analyses conducted in virtual environments. The conditions for visual and haptic exploration have advanced and so we can state that the virtual environment is fundamentally different when we experiment with interfaces of different generations. The creation of illusion effects specific to each system makes it possible to assume that the exploratory conditions are not similar and that comparison and transfer of information is difficult.

Applied research conducted on the processing of spatial information in a virtual environment sometimes conveys conflicting data in the learning domain. The exploration conditions cross-refer to two different types of space:

a. large action spaces (representation and orientation)
b. spaces with restricted action (detection and manipulation of simple objects).

This work shows that:

- the choice of reference markers is important for constructing a visual space which becomes the medium for spatial representation for the subject; indeed, a badly monitored activity could affect the subject's representation capabilities;
- the performance is sensitive to spatial distortion, restriction of the visual field and to the effects of depth.

The perceptive conflicts between movement and vision hamper the precision of the gesture and may modify the speed of movement of the gesture (Coello, Decety, Leifflen & Orliaguet, 1996) and, because of this, the concept of learning transfer between a virtual and a real world is compromised. On the other hand, the individual may acquire a performance on a particular sensorial capability.

The integrity of cross-reference in the virtual environment is an important factor. Is the perception of information received in a real environment faithful to the perception of information received in a virtual world? This concept of environmental fidelity involves the psychological judgement of the subjects plus the technological concept.

A virtual environment which makes it possible to lift the mask over the hidden task offers the subject the opportunity of memorising more complete information and enriching his spatial representation by the effects of 2D and/or 3D visualisation and of “rejects”; the hypothesis that the subject is able to transfer this information by limiting the affects of spatial knowledge may then be raised.

The virtual environment makes it possible to substitute one sensorial factor for another on the concept of amodality and in this way to isolate a sensorial process in order to obtain a better understanding of it. These artefacts can also make it possible to limit the mental load on the subject facing a heavy use of the equipment.

There are numerous controversies on:

- whether there is a need to reproduce identically the needs felt in a real environment for a simulation tool when there is a risk of its resulting in a heavy mental loading on the subject;
- the employment of cognitive capabilities in co-ordinating information from visual space and motor space in the virtual environment;
- the development over time of processes used in virtual environments;
- the need for modulation for subjects’ interpersonal dimensions during information processing in virtual reality.

6. The Results of Experiments in Compared Environments

The objective of the thesis was to compare conditions for exploring a shape under real and virtual environments using simple interfaces for assessing subjects’ performance on the acquisition of spatial information.

6.1 The needs of subjects in the real environment

The three shapes proposed require the operator to make a double category recognition: object is round or angular; if angular decide the aperture angle. These are exocentric factors that the subject will seek to identify in order to

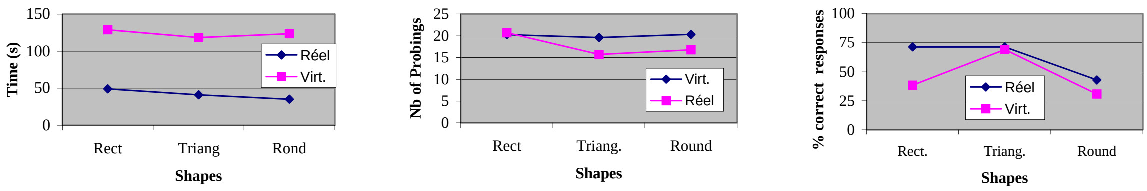
differentiate the three shapes. This observation has facilitated the breakdown of the gestures into these different translational and rotational movements.
We then observe that, in order to detect a shape:

- performance is not linked to expertise, which enables us to formulate the hypothesis that the subjects' personal strategies would have a discriminatory nature independent of the training received;
- the subjects' performance varies as a function of the simple shapes to be identified;
- the subjects' performance varies as a function of the environments concealing the shapes;
- processing the information for detecting a shape and its texture could be first class in the various uni-sensorial or bi-sensorial modalities while revealing a graduation in performance measurement;
- the cognitive style of dependence and independence with regard to the visual field is insufficient to explain the subject's strategy when he is referring mainly to the visual field;
- individuals have identification strategies for breaking down a shape according to the shape strategy concepts.

6.2 Comparison of learning in real and virtual environments when making a detailed gestural manipulation remotely

The results were obtained from variance analyses in order to measure four main effects which were: the environmental condition; learning; visual modality; and interpersonal variability.

6.2.1 The effect of the environmental condition enabled us to distinguish between real and virtual performance



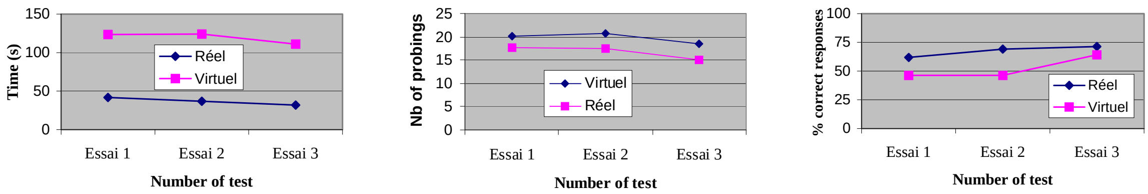
a: Probing time **b:** Number of probings **c:** Percentage of correct responses

Figure 1: Probing times, number of probings and percentage of correct responses of shapes for 1 test as a function of environment conditions.

The detection of a shape in the virtual environment required three times as long as in the real environment. The average time in the real environment was 42 s whereas it was 123 s in the virtual environment. Whereas the number of probings remained similar, the mean deviation between the two situations was three points. In contrast, the quality of the responses varied between the two situations. In the real environment we had 62% correct responses whereas in the virtual environment we obtained 46% correct responses (average for the different shapes).
The shape influences the subject's performance. The triangle, with a quick detection time, had the best success percentage in both conditions. The rectangle and the round shape showed detection conditions were more difficult, increasingly so in the virtual environment both for time and for correctness of response (Figure 1).
In order to determine a geometric shape the probing time increased considerably in the virtual environment, but, however, without, the action undertaken by the subject being changed significantly, and without this increase in time affecting the quality of the response. Although there was no increase in motor activity (that is in the number of probings), we did note that the time spent on concentration, attention or reflection was longer for an achievement of detection capability which was less difficult than one in a real environment. Although the subjects spent far more time to achieve an identical result, we can reckon that this difference in time is marked by the modification of sensori-motor activity and/or cognitive activity in order to compensate for the search for new identifying markers.

6.2.2 The learning effect

Learning was measured by systematic repetition of three tests for all subjects. This enabled us to eliminate the effect of chance.
Here, we found that the deviations in detection time between the first and third test were significant (Figure 2). For the first test we recorded 123s and for the third, 111 s. Differentiating the two environments we observe that the time curves decrease in parallel with the tests (Figure 1).
In terms of the quality of the response, the percentages increased during the three tests. They varied from 54.5% in test 1 to 66% in test 3, and the percentage of correct responses in the virtual environment showed a greater increase between test 2 and test 3.
Once again, for the number of probings we identified a slight deviation between test 1 (17.5%) and test 3 (15.5) which is not significant. Repeating the tests made it possible to commence learning under both environmental conditions (Figure 3).



a: Detection time **b:** Number of probings **c:** Percentage of correct responses

Figure 2: Detection times, number of probings and percentage of correct responses of shapes for the three tests in real and virtual environments.

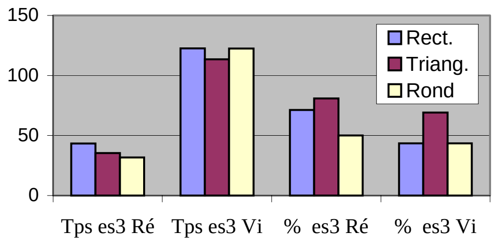


Figure 3: In the learning situation, comparison of time and % of correct responses for shapes between real (Re) and virtual (Vi) for test 3 (T3).

This learning is identified in real and virtual environments. Although the changes in time and number of probings are small we observe there is a marked increase in the percentage success. The display of the third test results also shows disparities in the detection of shapes.

6.2.3 The effect of visual modality in the learning situation

The contribution from this visual modality is measured on a special aid provided by the virtual environment. After each test, half of the subjects received a display of the probing points that they were able to superimpose over the shapes sought.

The sub-group benefiting from the visual aid detected shapes more quickly (99 s) than the sub-group without the aid (119 s), still with a deviation of 3 points on the number of probings. In contrast, with quicker detection the subjects using the aid answered with 70% correct responses whereas the sub-group without the visual aid had only 50% correct responses.

The presentation of visual information was envisaged as an aid to detection. This visual window offered space for reflection, allowing the subject to readjust the processing of the information obtained blind in order to confirm or reject the shape detection decision.

It appears that the subject reinforces his action at each new test in the virtual environment by still mobilising his sensori-motor activity just as much. In contrast, we found a slight reduction in the time, in parallel with an increase in the quality of the responses. The visual aid became an aid for representing the shape. It caused reflection on the action and enabled the subject to readjust his strategy for the next test.

6.2.4 The personal effect

This effect was measured initially after the subjects were divided into 4 sub-groups in accordance with the GEFT (group embedded figures test) which is a perceptive test measuring the capability of subjects to extract a simple shape from a complex figure.

On a test and per sub-group, the average detection times ranged from 68.62s to 112.36s for the number of probings from 50.5 to 82, and for percentages of correct responses from 47.5 to 80.

The variation in data between sub-groups did not correspond to the results expected. The performance specified by sub-group 1, identified as dependent with regard to the field, distinguished processing strategies defined in reference to Huteau's concept. In contrast, sub-group 4, categorised as independent with regard to the field, represented an "economic" processing strategy for the three variables: time; number of probings; and percentage success. These data are currently being studied to break down the perception strategies.

7. Subjects' Perception of the Virtual Environment: Review of Conversations

Subjects' conversations during the experiments provided us with the following information.

7.1 Visual perception of the context

The visual perception of an object in our case revealed a dispersion on the size of the object which was assessed at between 5 and 30 centimetres (the true dimensions being 15×15). Evaluation of the size of an object in a virtual environment was often unrealistic and varied from person to person with some over- or under-estimating, but others correctly assessing the dimensions of the target.

The visual perceptive conflict is the search for a visual compromise between the intention of accurately aiming a probing point and the possibility of reaching this precise probing point. In the present case, the subjects had to seek to align several probing points in order to create visual markers of the shape and to trace a curve, a straight line or an angle.

All subjects mentioned visual fatigue after the repetition of three tests interspersed by returns to a real environment. It was for this reason that we limited the learning to three tests.

7.2 Sensori-motor perception of the context

The time taken to perform a task under motor control.

The relationship of speed/accuracy of arm and hand movements, measured by the time between picking up the probe in the hand and identifying the target, was clearly modified in the virtual environment. For the detailed and precise gestural movements of the mine clearance expert, the subject will have to slow down his movement by continual monitoring in order to adjust his gesture. This means that the subject is going to have to adapt his sensori-motor movement by developing a slower gestural movement in order to achieve success or otherwise of the task undertaken.

The lack of sensori-motor and haptic information.

The virtual presentation of the target (visual and auditory factors) lacks haptic information, located mainly on the edges of the target. The need for this is revealed in the mine clearance expert's gestures by the manner of proceeding to obtain accuracy for the angular or rounded criterion.

8. Conclusion

For this paper, comparisons of conduct and learning in two environments enable us to report that the performance acquired when making precise gestural movements in a virtual environment is lower than the performance achieved in a real environment. However, we can state that repeating the tests enhances the speed/accuracy factor. The improvement seen in the two situations demonstrates that the subjects adapt and develop with this new environment. One advantage of the virtual environment in learning compared with the real environment is that it offers the possibility of calibrating the task on one or more modalities in order to measure the significant individual and collective performance on simple tasks. It could become a simulation tool of benefit to mankind, offering the possibility of isolating or combining several types of information in order to verify the specific needs of the individual.

9. Development

The results of this study have enabled us to specify the changes to the virtual environment needed for the mine clearance expert's DRI task. The new application uses:

- a Proview 60 helmet which, combined with a more powerful machine and graphics cards, has made it possible to improve the visual aspect and to stabilise the image;
- modelling of two real mines;
- a Phantom 1.5 3DoF force feedback arm which makes it possible to provide haptic effects (especially when detecting collisions), a more precise manipulation of the probe position (which enables the detail of gestural movement to be increased).

For budgetary reasons when specifying these changes, the force feedback was limited to translational movements. While it is necessary, theoretically, to constrain the probe to 5 degrees of freedom (3 translations and 2 rotations — pitch and yaw) for guiding the probe over the ground, this is not possible with the 1.5 3DoF version of the Phantom. In the absence of guidance, the displacement of the point of intersection of probe with the surface of the terrain, due to lack of gestural accuracy, is visualised in the virtual world and leads to a perceptive conflict.

In order to overcome this technology limitation, as soon as the end of the probe contacts the ground it is subject to guidance by a point within a tube along the probe axis of incidence, and the image of the probe is locked to this axis.

This artefact serves as a decoy for the operator's sense which, when the probe is free to move in rotation, works along a single translational axis.

In version 2 (addition of force feedback) this demonstrator could become a genuine tool for learning mine clearance strategy, enabling the instructor to validate the relevance of probings (searching for limits, width and height, detecting contours, enabling the shape to be identified), minimising the number of probings by developing strategies depending on the sensorial data received and thereby increasing the reliability of decisions in a real operation. In time, this tool could also make it possible to teach the technique to civilian populations and thus accelerate the decontamination process which is still long, costly in terms of money and also of human life.

Technology development already permits us to envisage version 3, a portable system which, by mathematical analysis of the probing geometry and comparison with a mine database, can offer a genuinely improved aid to decision making and processing in real operations. The greatest problem is to obtain a system which is not liable to trigger the mine irrespective of the latter's technology and therefore this means a system which does not emit a signal or signature of any kind.

With sociological problems overcome, we can envisage using a master force feedback arm to remotely operate a slave arm fitted with a probe; while retaining the skill aspect of the sapper's job, it would then become possible to shift the task towards the rear and thereby make mine clearance operations safer.

It nevertheless remains true that, beyond technology, the best way of obtaining terrain completely free from the presence of mines is not to mine it in the first place.

This page has been deliberately left blank

Page intentionnellement blanche

Acquiring Real World Spatial Skills in a Virtual World

Bob G. Witmer, Bruce W. Knerr and Wallace J. Sadowski Jr.
US Army Research Institute[1]
12350 Research Parkway
Orlando, FL 32826-3276
USA

Summary
In rehearsing specific missions, soldiers frequently must learn about spaces to which they have no direct access. Virtual Environments (VE) representing those spaces can be constructed and used to rehearse the missions, but how do we ensure their effectiveness? The US Army Research Institute was among the first to demonstrate that spatial knowledge acquired in a virtual model of a building transferred to the real world. While route knowledge was readily acquired in a VE, configuration knowledge (distance and direction to locations not in the line-of-sight) was not. Spatial learning in the VE was hampered not only by disorientation resulting from a narrow FOV and multiple collisions with walls, but also by participants' inability to accurately estimate distances in VEs. Poor distance estimation in VE was linked to the reduced VE FOV and to verbal report procedures for making the estimates. Some improvement in distance estimates was obtained by adding auditory compensatory cues for distance and by using the non-visually locomotion technique for obtaining distance estimates. Armed with knowledge that some VE characteristics adversely affect distance estimation and configuration learning, we conducted research to determine if unique capabilities of VEs could compensate for those characteristics. We developed three VE navigation training aids**:** local and global orientation cues, aerial views, and division of the VE into distinctive themed quadrants. The aids were not provided when testing configuration knowledge. Training included a guided tour, free exploration of the VE and searching for designated rooms. Configuration knowledge tests included a shortest route test, a pointing task, and a map construction task. An aerial view was the most effective navigation aid, though its effectiveness depended on how it was used. Those participants who used aerial views to organize the VE and learn its layout during free exploration performed quite well, while participants who used it as a crutch to locate a particular destination performed worse than those without an aerial view. To ensure that VEs train effectively, we must recognize VEs' deficiencies, compensate for deficiencies whenever possible, and exploit VEs' unique training capabilities.

Introduction
The U.S. Army has invested heavily in the use of virtual environments (VE) to train combat forces, to evaluate new systems and operational concepts, and to rehearse specific missions. While the Army has focused mainly on simulations for mounted combat, there is also a need to train infantry and other dismounted soldiers. In training dismounted soldiers there are occasions (e.g., rehearsing a hostage rescue mission) in which the soldiers must learn about strategically important spaces to which they have no immediate access. Virtual environments can be constructed as a substitute for these spaces, but how effective are they? This paper describes a series of experiments that investigated the limitations of using VE for training spatial knowledge and how VE might be improved to meet Army human performance goals.

Although VE technologies such as helmet-mounted visual displays, head trackers, 3-D sound systems, haptic devices, and powerful graphics image generators have the potential to immerse dismounted soldiers directly in virtual training environments, their capability to provide effective training has yet to be ascertained. The effective use of VE for training requires more than just VE hardware and software. It also requires a body of knowledge that identifies the characteristics of VE systems that are required to provide effective training and the training strategies and features that are most appropriate for use with VE. In order to develop this body of knowledge, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) Simulator Systems Research Unit, initiated a program of experimentation to investigate the use of VE technology to train dismounted soldiers in 1992.

Experiment 1: Transfer of Spatial Knowledge
We were among the first to conduct research demonstrating transfer of spatial knowledge from VE to a real world environment (Witmer, Bailey, Knerr, & Parsons, 1996). For this research, a detailed model of a large office building was constructed using Multigen and World Tool Kit. The model was rendered using a Silicon Graphics Crimson Reality Engine and displayed via a Fake Space Lab Boom. The Boom consists of a high-resolution binocular display on the end of an arm that allowed six degree-of-freedom movement and thumb buttons for controlling forward and backward motion.

[1] For correspondence with author: Bob_Witmer@stricom.army.mil.

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

The participants were sixty college students who had no previous exposure to the building. Participants first studied route directions and photographs of landmarks, either with or without a map, then were assigned to one of three rehearsal groups. These were (1) a VE group that rehearsed in the building model, (2) a building rehearsal group that rehearsed in the actual building, and (3) a symbolic rehearsal group that relied on verbal rehearsal of the route directions. Participants were then tested in the real world building for transfer of route training.

Differences in training transfer were evaluated using a MANOVA with rehearsal mode, map, and gender as the independent measures. Only the main effect for rehearsal mode was significant ($p<.001$). A follow-up ANOVA indicated that this effect was significant for each of the dependent measures: route traversal time ($p<.001$); number of wrong turns ($p<.001$); and total distance traveled ($p<.05$). Participants trained in the building made fewer wrong turns ($t=3.25$, $p<.005$) and traveled less distance ($t=2.9$, $p<.01$) than did participants who were trained in the virtual environment (VE). VE participants, in turn, made fewer wrong turns ($t=-4.77$, $p<.001$) and took less time to traverse the route ($t=-5.82$, $p<.001$) than those who were trained symbolically.

In practicing the route, participants were expected to acquire some knowledge about the overall layout of building (i.e., the building configuration). Configuration knowledge was measured using the projective convergence technique (Siegel, 1981; Kirasic, Allen, & Siegel, 1984) and by measuring the capability of subjects to exit the building quickly using an unrehearsed route. The projective convergence technique requires participants to estimate the distance and direction to target locations not in the line of sight, and uses these estimates to determine the participant's perceived target location. The participants either draw lines to indicate the distance and direction to targets (in a non-immersive mode) or point to indicate bearing and verbally report their distance judgments in standard or metric units (in an immersive mode). Errors in estimated bearing and distance using this method may either be due to poor distance estimation skills or disorientation and a lack of knowledge regarding the designated target location. Hence it is not a pure distance estimation measure. MANOVA was used to assess differences in the amount of configuration knowledge. Surprisingly, there were no significant differences among the various rehearsal conditions ($p=.135$) and no significant differences as a function of map use ($p=.688$). Only the effect of gender was significant, with males performing better than females ($p=.015$). No significant interactions were found.

The results suggest that individuals can learn how to navigate a real world route by training in a virtual environment. While the VE used in this experiment was not as effective in training subjects as the actual building, it was much better than verbally rehearsing route directions, even for subjects who had previously studied a map. The effectiveness of the VE for acquiring route knowledge was probably limited by the display reduced field of view and by disorientation after collisions with virtual objects. These factors along with an unnatural interface that controlled movement through the VE. These factors along with participants' inability to judge distance in VEs may also have adversely affected the acquisition of configuration knowledge.

Experiments 2–5: Judging distance in Ves

To better understand why participants were unable to accurately judge distance in the VE, ARI investigators conducted a series basic research experiments in the area (Kline & Witmer, 1996; Witmer & Kline, 1998; Witmer & Sadowksi ,1998). Kline & Witmer (1996) and Witmer and Kline (1998) used magnitude estimation to measure participants' ability to estimate distances in a VE. The task was performed in a virtual office corridor with various floor and wall patterns and textures. Participants first estimated the distance to a standard stimulus (e.g., a cylinder at 100 feet)[2]. They received no feedback regarding the accuracy of their distance estimates to the standard stimulus, but were told that all subsequent estimates should be made relative to that standard. Actual distances varied from 1 to 12 feet in one experiment (Kline & Witmer, 1996), from 10 to 110 feet in another, and from 10 to 280 feet in a third (Witmer & Kline, 1998). The basic measure for all of these experiments, with the exception of Witmer and Sadowski (1998), was the reported target distance in feet or meters. The amount of error in these estimates was calculated as the difference between the estimated and true distance divided by the true distance. This error measurement is called relative error because it is the amount of error relative to the true target distance.

Kline & Witmer (1996) investigated how accurately stationary observers could estimate distance to a wall in a VE as FOV, texture, and pattern were varied. The observer's view was fixed (i.e., no head tracking}. The distances being judged were between 1 and 12 feet. The results indicated that a wider FOV (140H × 90V degrees) produced more accurate estimates than a narrow FOV (60H × 38Vdegrees), $\underline{F}(2,23)=5.85$, $\underline{p}<.01$. Distances were typically underestimated with the wide FOV and overestimated using the narrow FOV. For example, a target placed 5 feet from the observer was judged to be at 2.68 feet with the wide FOV and 8.73 feet with the narrow FOV. Significant two-way interactions of distance with texture, $\underline{F}(44,1054)=2.53$, $\underline{p}<.001$, pattern, $\underline{F}(22,3)=14.1$, $\underline{p}<.05$ and FOV, $\underline{F}(44,1054)=2.5$, $\underline{p}<.001$. indicated that these variables affected depth perception only at the shorter distances.

[2] Note: All distances are given in feet. Multiply by .3048 to convert to meters.

In another experiment, Witmer & Kline (1998) investigated the effects of floor texture and pattern on distance judgements to a cylinder for distances up to 110 feet. The observers were stationary and had a fixed view of the target scene (i.e., no head tracking). Participants grossly underestimated the target distance; the estimates averaged about 50% of the true target distance. This compares to estimates of approximately 75% of the true distance in a comparable real world environment. Cylinder size, $F(1,22)=38.67$, $p<.001$, distance, $F(5,18)=5.87$, $p<.01$, and the interaction of cylinder size and distance, $F(5,18)=3.97$, $p<.05$, significantly affected the magnitude of the VE estimates. The estimates were more accurate for the small cylinder than for the large cylinder. For example, a target placed 50 feet from the observer was judged to be 22.57 feet for the small cylinder and 18.91 feet for the large cylinder. Floor texture did not significantly affect either the distance estimates or the magnitude of the relative errors.

Witmer & Kline (1998) also reported the results of an experiment in which moving observers judged distance traversed for distances up to 280 feet. Half of the participants received compensatory cues (an audible tone every 10 feet) to help them calibrate their distance judgements to the true target distances. Although these cues were provided on only half of the trials, they improved performance to levels approaching perfect performance, $F(1,60)=11.49$, $p<.001$. The judgments averaged 96% of the true target distance when compensatory cues were present but only 67% of the target distance when compensatory cues were absent. The mode of locomotion used in moving through the VE (treadmill, joystick, or teleport) did not significantly influence the accuracy of the distance estimates, but speed of movement had a significant impact on estimation accuracy, $F(1,60)=36.15$, $p<.001$. Distance judgments were more accurate at the slow speed than at the fast speed. For example, a distance of 280 feet was judged to be 267 feet on the average when moving at the slow speed and 241 feet when moving at the fast speed. Accuracy of the distance estimates generally decreased as distance to the target increased, $F(7,54)=482.53$, $p<.001$.

The extremely poor VE distance estimates made by a stationary observer and the lack of substantial improvement in the accuracy of the estimates when observer movement was added (Witmer & Kline, 1998) suggests that either verbal estimates of distance are not very accurate or that VEs degrade distance estimation to a large degree. The ability of participants to accurately report distances in feet or meters varies widely among participants, and may be independent of their perception of target distance. These individual differences may inflate the amount of error observed in estimating target distance. To determine how much of the problem is due to the requirement to provide verbal estimates of distance and how much is due to VE factors, Witmer & Sadowski (1998) used non-visually guided locomotion (NVGL) to obtain distance judgements in VE and real world environments. Participants viewed a target for 10 seconds from a stationary position, forming a mental image of the target's location. They were then blindfolded and asked to walk to the target's location, keeping the target's location in their minds as they approached it and stopping when they thought they had reached it. They were asked not to count steps or time mentally. The distance judgments were performed both in a real world officer corridor and in a virtual office corridor modeled to simulate the real world corridor. The target, a construction cone, was clearly visible and distinct from the background at all distances. Participants made judgements for targets placed at distances between 15 and 105 feet. The distance judgements averaged about 85% of the true target distance in the VE and 92% of the true target distance in the real world environment. The differences between the distance judgements in the VE and in the real world were significant, however, $F(1,20)=4.41$, $p<.01$. The magnitude of the errors in the VE was nearly twice those obtained in the real world.

Implications of the learning transfer and distance estimation experiments

Our initial investigation of configuration learning (Witmer et al., 1996) suggested that distance estimates in VE were poor. Witmer and Kline (1998) confirmed this, showing that distance estimation in a VE is significantly less accurate than in the real world. Kline & Witmer (1996) demonstrated that reducing the FOV for one of the devices (BOOM2C) could affect not only the amount of error in distance estimates, but also the direction of that error (underestimates vs. overestimates). The hypothesized that narrow FOV produced less accurate estimates by reducing or eliminating linear perspective cues. Witmer & Kline (1998) found that manipulation of textures did little to eliminate the observed deficits in performance. Although target size did influence performance, manipulation of the size of unfamiliar objects is not a practical solution. Taken together, these studies suggest that VEs distort monocular or stereoscopic distance cues, negatively impacting the distance judgements in those VEs.

We had anticipated that providing the cues for distance associated with movement would compensate for the distortion of other distance cues in VE, resulting in substantial improvements in performance. However, Witmer & Kline (1998) found that neither movement method nor edge rate markedly changed the distance judgments. These results indicate that proprioceptive cues and visual flow cues may not play a major role in making distance judgements in a VE. In contrast, movement speed clearly influenced distance judgments, suggesting that the time spent covering a distance changes one's perception of distance traveled. This research also suggested that distance perception in VE could be recalibrated cognitively by providing compensatory cues for distance. This cognitive recalibration may or may not extend to other distances or

to other environments, however. Witmer & Kline (1998) did not collect data that would answer questions about transfer of estimating skill to other distances or environments.

Using NVGL to evaluate the accuracy of VE distance estimates altered our working hypothesis regarding how much VE degrades distance estimates. This procedure yielded more accurate VE distance estimates, suggesting that the use of verbal distance estimates is partly responsible for the poor performance observed in our research. However the magnitude of the errors in VE using the NVGL procedure was still twice that observed in the real world, establishing beyond any reasonable doubt that VEs are distorting perceptual judgments of distance.

Factors influencing VE distance judgements

What factors might be responsible for this distortion? In our search for an explanation it is important to remember that the performance decrements were found across various VEs using different display devices, and with varying movement conditions. It is also important to keep in mind the distances investigated in each experiment, because the effective range of various distance cues vary with the distance being judged.

To understand why VE distorts distance perception at the target distances investigated, we need to know which distance cues are effective at those distances, and to assess the extent to which these cues were present or absent in our research. Cutting and Vishton (1995) have identified which depth cues are most effective at different distances and related these cues to three egocentric regions or zones of space: (1) personal space extends just beyond arms reach and refers to space used by a static observer; (2) action space extends to about 100 feet and refers and includes distances in which an observer can throw an object to another person or easily talk to others; and (3) vista space extends beyond 100 feet. Kline & Witmer (1996) studied both personal and action space. In personal space the most important depth cues are occlusion, binocular disparity, relative size, convergence and accommodation. The remaining studies investigated action space and vista space. The primary distance cues in action space and vista space are the pictorial cues, including occlusion, height in the visual field, convergent linear perspective, relative size, and relative textural density. In addition, two other distance cues, binocular disparity and motion perspective are effective distance cues in action space. Note that accommodation and convergence are not effective depth cues in action space or vista space.

Witmer & Kline (1997) have shown that while relative textural density influences distance estimates in VE, its effects are typically too small to account for the differences between real world and VE distance estimation performance. Similarly adding observer movement, which provides motion perspective and other movement related cues does not eliminate the deficits in performance in VEs (Witmer & Kline, 1998). Research by Wright (1995) and Witmer & Kline (1996) suggests that simply using a high resolution or wide FOV VE display cannot erase the deficits in perceived distance. Although occlusion is probably the most powerful depth cue in action space, it was not a factor in our distance estimation tasks. Of the remaining distance cues listed by Cutting & Vishton (1995), height in the visual field, convergent linear perspective, relative size, and binocular disparity appear to be the most likely candidates for explaining the observed discrepancies between VE and real world judgements of distance.

The National Research Council (1997) has suggested that the restricted FOV provided by VE displays must degrade height in the visual field and convergent linear perspective as cues for distance at some point. The limited vertical FOV found in most VE displays (ranging from 40 to 90 degrees) may be responsible for this degradation. By comparison, the real world vertical FOV is approximately 120 degrees. A reduced vertical FOV may result in distant objects appearing closer in VE than they would in the real world because these objects would be compressed into a smaller visual frame as they recede into the distance. Kline & Witmer (1996) showed that a reduced horizontal FOV could also adversely impact the accuracy of distance estimates by reducing or eliminating linear perspective cues. Because linear perspective cues are among the most effective distance cues in simulated environments (Surdick et al., 1997), reducing or eliminating these cues can have a major impact on the accuracy of distance estimates.

In VEs, emulation of binocular disparity is achieved by presenting different images to the two eyes with some central area overlap. While this technique may provide the illusion of depth in VE, it may not faithfully reproduce real world depth. Cutting & Vishton (1995) noted that early stereoscopic pictures enhanced the distance between the eyes to show large expanses and cityscapes, diminishing the effective size of the objects seen. Relative size may be important factor at the closer distances because the perceived size of an object accelerates as the distance to the object decreases, yielding a looming effect. Accommodation and convergence cues are not accurate in VEs, a fact that researchers often use to explain poor distance estimation in VEs. However, these cues are only important for judgments in personal space and at the shorter distances within action space.

Additional research is needed to determine which of the distance cues operating in action space are most responsible for degrading distance judgements in VE. Once the causes of this degradation are isolated, we can begin working toward a solution. The solution may be as simple as increasing the VE display vertical or horizontal FOV, or adjusting the overlap in VE stereoscopic viewing devices. On the other hand, it may involve major technological advances, such as inventing new

techniques for emulating binocular disparity in VE displays.

Having identified some of the factors that affect distance judgements in VE, we turned our attention back to how to best use VEs for training configuration knowledge. Our approach was to utilize unique capabilities of VE that might compensate for its inherent deficiencies (e.g., VE's tendency to distort distance judgements).

Enhanced VEs for spatial knowledge acquisition
A computer model of one floor of a large office building, used in previous research (Bailey & Witmer, 1994; Witmer et al., 1996) was adapted for this experiment. All passageways in the virtual building were widened to reduce collisions, an improved collision detection algorithm was introduced that decreased the need to back away from objects following a collision, and additional rooms were modeled. Separate VE models were constructed to represent the standard and enhanced environments. The enhanced environment was created by adding theme objects and sounds to the standard environment model. The models were created using Multigen II software and rendered by a Silicon Graphics Onyx with eight 200MHz processors and three RealityEngine2 Graphics Pipes. Both models were displayed using a Virtual Research V8 Helmet-Mounted Display (HMD). Locomotion through the VE was achieved by virtual walking in the safety pod shown in Figure 1. Head and body movements were independently tracked.

Figure 1: Safety Pod for Virtual Walking

The participants were sixty-four college students who had no previous exposure to the building. Following a brief train-up, the participants were randomly assigned to one of eight treatment groups, who received different levels of navigation aids. Depending on group assignment, a participant experienced either the standard or enhanced VE, received orientation cues or did not, and could chose to view the VE from an aerial perspective or was restricted to viewing the VE from the normal perspective. Orientation cues included an arrow projecting from the chest of the participant's avatar and a flagpole visible throughout the environment.

Groups having an aerial perspective could view the VE from heights of 49, 98, and 394 feet for a period of up to one minute. After one minute, they automatically returned to the normal perspective view. The viewing heights were selected such that participants could see either the whole third floor layout at once at 394 feet or parts of the layout at 39 and 98 feet. More objects in the environment could be recognized at the lower viewing heights. Figure 2 shows the VE from a viewing height of 98 feet. While in the aerial mode participants could further explore the environment by flying to other aerial locations (accomplished by walking in place). To return to ground level they pressed the thumb button on their hand controller, and gradually descended to reenter their virtual body at the exact location where they left it when they started to fly.

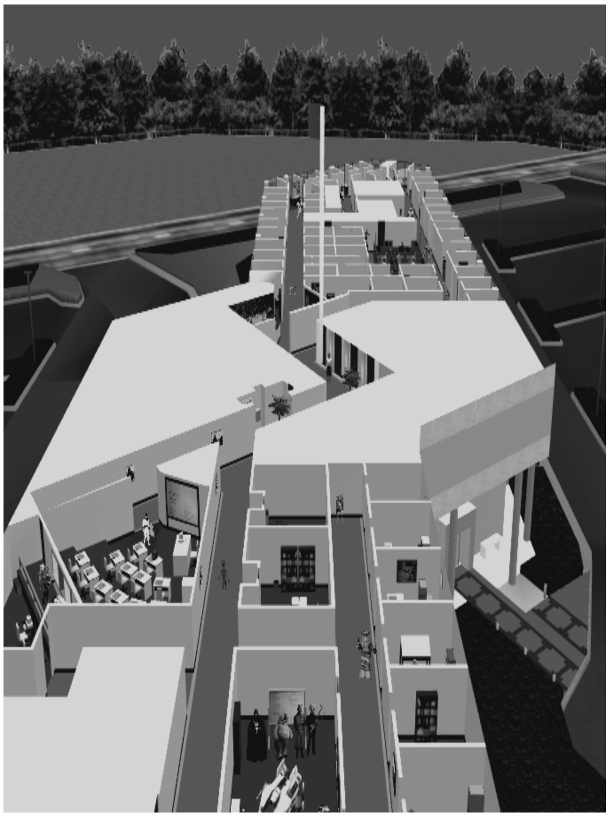


Figure 2: Aerial View of Third Floor Viewed at 98 feet

The enhanced environment model was divided into four themed quadrants or districts. Groups exposed to the themed environment encountered sights and sounds associated with the themed quadrants. Each destination had a memorable theme object located inside the room and an associated sound that became louder as the participant approached the destination room. Additional theme objects were positioned along the building

corridors, but no sounds were associated with these additional objects. The themes embedded in the quadrants were a tropical island theme, a wild animals theme, an extraterrestrial (or outer space) theme, and a sports theme. Upon encountering a theme object located inside one of the destination rooms, participants were asked to identify the theme represented by that object. This encouraged participants to associate destination rooms with their location in a particular quadrant.

The orientation cue groups were asked to relate their current position to their starting position marked by a virtual flagpole. This was accomplished by facing the flagpole upon reaching each destination. The flagpole served as a global orientation cue that allowed participants to continually update their current position based on their known starting position. Participants were told to use the arrow projecting from the chest of their avatar as an indication of their current heading and as a way of aligning their virtual body so as to avoid collisions with walls and doorways.

Individual training and testing phases comprised the research. During the first training phase participants followed a virtual tour guide through the VE, pausing at each destination room, and identifying it by name. The tour guide verbally described the 'non-theme related' distinguishing features of each destination. In the second training phase, participants explored the VE freely, while trying to locate and identify each previously visited destination. In the final training phase, participants attempted to take the shortest route from the third floor lobby to each named destination. If the participants did not find the destination within three minutes, they were verbally guided to it. Knowledge of the building configuration was tested by asking participants to complete the following tasks: (1) take the shortest route between designated rooms, (2) estimate the distance and direction to locations not in the line-of-sight, and (3) place room cutouts in their correct locations on a map. Similar to the NVGL procedure, participants estimated distance by walking the straight-line distance between their current location and the perceived location of the destination without vision. Navigation aids were not provided during the testing phase. A follow-up room placement test was given one week after the initial test to examine retention of configuration knowledge.

The purpose of the navigation aids was to offset the effects of VE deficiencies that interfere with the acquisition of configuration knowledge in a VE. The orientation cues had no significant effects on configuration knowledge acquisition, $\underline{F}(4,51)=2.05$, $\underline{p}=.10$. Participants receiving the enhanced environment performed better during training than those who received the standard environment, $\underline{F}(4,51)=2.80$, $\underline{p}<.05$, but not on the tests of configuration knowledge. Only the participants who received an aerial perspective view performed significantly better both during training, $\underline{F}(4,51)=5.69$, $\underline{p}<.001$, and on the configuration knowledge tests, $\underline{F}(6,50)=3.44$, $\underline{p}<.01$. Participants with an aerial view during training also performed better on the 1-week retention test, $\underline{F}(1,51)=9.76$, $\underline{p}<.01$.

The effectiveness of the navigation aids, including the aerial view, seemed to depend on how the participants used the aids. When the aids were used as a crutch to quickly find a room, they were not effective. Similarly in those cases where the navigation aids increased the workload beyond what the participants could handle, no performance gains were realized. The navigation aids seemed to work best when participants were able to use them to mentally structure the environment. For additional discussion of the effects of these navigation aids, see Witmer, Sadowski, and Finkelstein (in press).

Conclusions

What then must be done to ensure that training in virtual environments meets military human performance goals? The first step is identify the shortcomings of VE that adversely affect VE training effectiveness and link these shortcomings to specific performance deficiencies. For example, in spatial learning, a reduced FOV in VE was linked to poor distance estimation and spatial disorientation, ultimately impairing the acquisition of route and configuration knowledge. The next step is to determine if the deficiency can be addressed directly, or if not, how to compensate for the deficiency. Currently increasing the FOV for VE displays is an expensive proposition and large FOV devices may sacrifice resolution for the larger FOV. We used auditory cues to compensate for poor distance estimation in the VE and showed that the estimates were improved even when the cues were not present. We adopted the NVGL procedure to reduce the affects of individual differences on distance estimation tasks, and used it to measure distance in the projective convergence test. We took steps to reduce collisions in VE, thereby reducing the amount of disorientation that occurred with a narrow FOV display. We also increased the effective FOV by providing participants with an aerial view leading to improved acquisition of configuration knowledge. In searching for effective compensatory mechanisms, some promising factors had little practical effects. A more realistic walking interface (i.e., a treadmill) did not improve distance estimates and dividing the environment into themed quadrants or districts did not improve the performance on tests of configuration knowledge. This demonstrates the importance of evaluating VE interfaces and training enhancements in controlled experiments before implementing them in military training environments.

References

Bailey, J.H. & Witmer, B.G. (1994). Learning and transfer of spatial knowledge in a virtual environment. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society, 1158-1162.

Cutting, J.E. & Vishton, P.M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S.J. Rogers (Eds.), *Handbook of perception and cognition: Volume 5, Perception of space and motion* (pp. 69-117). New York: Academic Press.

Kirasic, K.C., Allen, G.L. & Siegel, A.W. (1984). Expression of configurational knowledge of large-scale environments: Students' performance of cognitive tasks. Environment and Behavior, 16 (6), 687-712.

Kline, P.B. & Witmer, B.G. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, 1112-1116.

National Research Council (1997). *Tactical display for soldiers: Human factors considerations.* Washington, DC: National Academy Press.

Siegel, A.W. (1981). The externalization of cognitive maps by children and adults: In search of ways to ask better questions. In L.S. Liben, A.H. Patterson & N. Newcombe (Eds.), *Spatial representation and behavior across the life span: Theory and application*. (pp. 167-194). New York: Academic Press, Inc.

Surdick, R.T., Davis, E., King, R.A. & Hodges, L.F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *PRESENCE: Teleoperators and Virtual Environments, 6 (5),* 513-531.

Witmer, B.G., Bailey, J.H., Knerr, B.W. & Parsons, K.C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human Computer Studies, 45,* 413-428.

Witmer, B.G. & Kline, P.B. (1997). *Training efficiently in virtual environments: Determinants of distance perception of stationary observers viewing stationary objects* (ARI Research Note 97-36). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Witmer, B.G. & Kline, P.B. (1998). Judging perceived and traversed distance in virtual environments. *PRESENCE: Teleoperators and Virtual Environments, 7 (2),* 144-167.

Witmer, B.G. & Sadowski Jr., W.J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human Factors, 40 (3),* 478-488.

Witmer, B.G., Sadowski Jr., W.J. & Finkelstein, N. (in press). *Training dismounted soldiers in virtual environments: Enhancing configuration learning* (Draft ARI Technical Report). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Wright, R.H. (1995). *Virtual reality psychophysics: Forward and lateral distance, height, and speed perceptions with a wide angle helmet display* (ARI Technical Report 1025). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

This page has been deliberately left blank

Page intentionnellement blanche

Advanced Air Defence Training Simulation System (AADTSS)

Virtual Reality is Reality in German Airforce Training

M. Reichert
Federal Office of Defence Technology and Procurement
FE I 4
Ferdinand-Sauerbruch-Str. 1
D-56073 Koblenz
Germany

This article describes the AADTSS simulation system and explains the reasons why it was realised with Virtual Reality technology.

The requirements

The programme started with the following main requirements:

- STINGER team training (commander, gunner)
- transportable, mobile
- size of scenarios: 360°×130°
- 8 targets, 8 effects, 2 missile firings at the same time
- long range aircraft detection and identification
- fast database generation system.

Why Virtual Reality?

As you can see, there are the contradictory requirements size of “scenario” and “transportable”.
These requirements make it impossible to use a normal dome-display-system. The solution is Virtual Reality.

The technical solution

The AADTSS simulator is integrated in a container.
Both, the commander and the gunner are wearing Head-Mounted-Displays (HMDs). Because of the requirement “long aircraft detection/identification”, the resolution per eye is 1280×1024 pixels. The HMDs are without a see-through-option, because of the better contrast and the advantage, that there is no need to switch of the lights in the container.

The two team members need to communicate acoustically and optically. Because of the closed HMDs the students cannot see each other. This problem was solved by modelling the commander and the gunner as avatares.

This solution might be a little bit funny, but it is well accepted by the soldiers.
The commander is tracked by an inertial tracking system, the gunner is tracked by an optical tracking system. Magnetic tracking systems are not suitable for use in environments like containers made of metal.
Orientation rings for the commander and the gunner are integrated in the container. This solution is necessary because of the HMDs without see-through-option.

Database generation system

The generation of databases is based on stereoscopic photos. It allows the generation of scenarios, targets and flight paths and is independent of the simulator. It consists of one workstation, the stereo-camera-system and one control-PC.
The main advantage of this system is, that there is no need for geographical data like maps or DTED and DFAD data.

Milestones

- 04/1995 First requirement
- 11/1997 Troop trial unit
- 05/1999 Final configuration

Paper presented at the RTO HFM Workshop on “What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?”, held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

This page has been deliberately left blank

Page intentionnellement blanche

"What is Essential for Virtual Reality to Meet Military Performance Goals" Performance Measurement in VR

Lt. Jim Patrey[1], Robert Breaux, Andrew Mead & Elizabeth Sheldon
Virtual Environment Training Technology (VETT) project
Naval Air Warfare Center Training Systems Division (NAWCTSD)
12350 Research Parkway
Orlando, FL 32826-3275
U.S.A.

One of the unique attributes and potentially greatest assets of virtual environments is the unique ability to comprehensively measure human performance. In the real environment, measuring human behaviors is usually, though not always, feasible and typically extremely effort intensive and cost-prohibitive. Similarly, there is substantial environmental variability that can have pervasive effects on human performance, but is beyond any feasible, economic data capture. Virtual environments instill the capability for comprehensively monitoring both user inputs and interactions and the environment (as well as control the virtual environment and thereby eliminating confounding variables with precision beyond that of real environment lab research).

Monitoring and measuring human behavior in this fashion provides three invaluable elements. Firstly, it furnishes a valuable research tool for the development of outcome measures for performing research. Secondly, performance measurement has training value for assessment and evaluation. The derivation of accurate performance measures can enable improved proficiency and reduced training time when implemented in a training curriculum. Finally, the development of performance measures can facilitate the development of intelligent tutoring systems and thereby cost-effective, stand-alone training systems. Measuring human performance can be of great use in the facilitation and maximization of training.

Performance measurement involves three distinct processes: Identification, Monitoring, & Evaluation. *Identification* is the determination of the significant measures of performance for a given task. This is typically accomplished via cognitive task analysis and intense subject-matter expert (SME) interviews and observation and/or statistical analytic techniques occurring after the observation of real world performance. These approaches are the two traditional approaches to performance measure development.

The advent of virtual environments has fostered the development of two new approaches to the development of performance measures - cognitive model driven and data-driven approaches. Cognitive models enable a new method of performance measurement. Through traditional approaches (such as SME interviews) a cognitive model can be developed for a given task (in truth, a cognitive task analysis is a variant of a cognitive model, typically represented in GOMS format). There are a host of cognitive modeling approaches (discussed in detail in Pew & Mavor, 1996), but they all generally afford identification of cognitive variables not easily discernable through traditional approaches. However, the usefulness of such models for performance measurement is dependent on the accuracy of the model and the development of cognitive models can be resource-intensive, particularly for complex tasks.

Data-driven approaches are also afforded by virtual environments. The ability to thoroughly monitor and record all actions and interactions in a virtual environment enables data mining approaches to provide value to the determination of performance measures. There are numerous data-driven techniques for mining data (such as neural networks, genetic algorithms, evolutionary computing, etc.), but it is fuzzy sets theory, or fuzzy logic, which may hold the most promise for identifying crucial aspects of human performance. Unlike other approaches, fuzzy logic preserves the semantic value of the input variables. Output from fuzzy models meaningfully represents human behavior and can be directly applied to performance measure development (Cowden, Burns, Casey, & Patrey, 2000).

It is likely that all of these approaches should be integrated to fully profit from virtual environments for the identification of performance measurement. Ideally, we will someday be able to place a SME in a VE to perform a task and have hybrid models (of both top-down cognitive models and bottom-up data-driven models) monitor the virtual world and generate performance models that produce measures of performance.

[1] For correspondence with author: patreyje@navair.navy.mil

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

VE-based performance measures cannot be developed without *Monitoring* the virtual environment. Accomplishing this requires monitoring behaviors and their consequences within the VE. Behaviors include active behaviors such as control inputs and verbal commands as well as passive behaviors such gaze surveys. The consequences of these actions include movement through the VE and interactions with and within the VE resultant from user behaviors. The principal behaviors and consequences must accurately represented, inherently measurable, and recorded for the effective use of VR for performance measurement.

Implicit in this is the indispensability of adequate modeling of the VE. All salient cues must be represented with suitable fidelity within the VE for the performance measures reaped to represent real world performance. This may be the greatest challenge for the practical use of VE for performance measurement. It generally behooves VE developers to minimize the fidelity in order to minimize processing demands and cost. The level of fidelity should be mapped to the task fidelity requirements so that 'training' fidelity, the level of fidelity required to meet training requirements, can be attained. The role of the SME cannot be underestimated in fulfilling this balance between minimal fidelity and requirements. Achieving this necessitates thorough front-end analysis prior to significant investment in development of the VE.

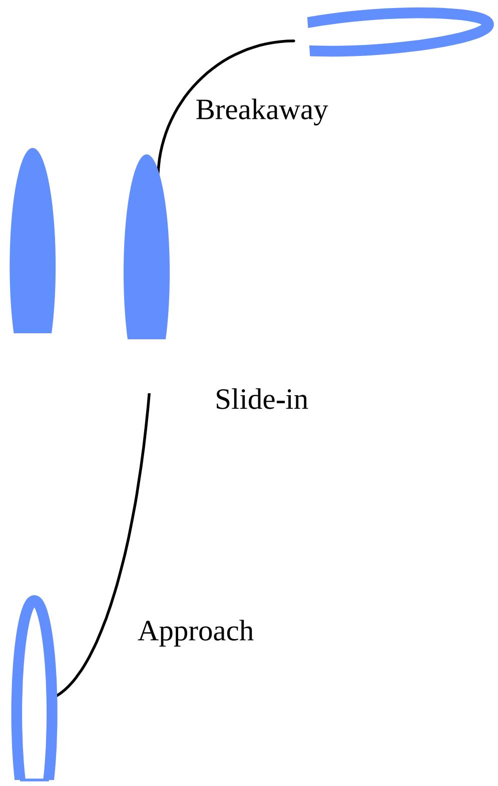
Finally, the effective use of VE in performance measurement should also provide performance *Evaluation*. Beyond identifying and monitoring performance measures is the need to discriminate good/expert performance from bad/novice performance. This is most meaningful for VE in the context of developing intelligent tutoring systems (ITS), but also permits structured, empirically based, objective feedback in any circumstance.

Derivation of evaluatory measures of performance (MOPs) is generally accomplished through methods similar to identifying performance measures. Traditional methods include SME ratings of performance (typically gathered through observation of another's performance) and statistical analysis. Cognitive model and data driven approaches also hold promise for evaluating performance (particularly in contrasting novices and experts), but they have not been applied as extensively in this domain.

Traditional performance measure development for virtual underway replenishment

An immersive virtual environment has been developed for underway replenishment (UNREP) with a U.S. Navy Cruiser (see Davidson, 1997 and Martin et al., 1998 for more information on the virtual UNREP). An UNREP involves the transfer of fuel, stores, ammunition, and people from one vessel to another while underway. It is comprised of four distinct phases (see Figure 1): 1) *Approach* - from awaiting station to bow-stern crossing (overtake oiler & attain lateral separation), 2) *Slide-in* - transition from approach to alongside (match velocity), 3) *Alongside* - stationkeeping (maintaining proper lateral separation and matched velocity), & 4) *Breakaway* - separation of own ship from oiler.

Figure 1. Depiction of the phases of Underway Replenishment.



The ship is controlled by the Conning Officer via verbal commands to a virtual helmsman. The verbal commands are broken down into two main types: control commands and requests for information. Control commands include <u>engine commands</u> such as all stop, all back, all ahead, indicate knots, increase turns, & decrease turns and <u>rudder commands</u> such as rudder amidships, steer course, left rudder, & right rudder. The Conning Officer can also make “requests for information” regarding rudder angle, relative bearing, true bearing, heading, speed, & range. These shiphandling behaviors provide a solid foundation upon which to develop MOPs.

Iterative inputs from SMEs also identified ship dynamic features indicative of good performance. These parameters vary depending upon the phase stage (approach, slide-in, alongside, or breakaway), but generally include relative positional data (vertical separation, lateral separation, & bearing) and relative velocity. The following depicts the statistical analyses conducted in pursuit of MOP identification.

Method

Subjects

Twenty-six (26) male Navy personnel (students & instructors) of the Surface Warfare Officer's School (SWOS) participated as subjects. Due to technical errors in the VE data collection process, data from eight (8) subjects, were not included in the analysis. The level of duties represented in the sample were Ensign (ENS, n = 6), Division Officer (DIVO, n = 4), Department Head (DH, n = 3), and Commanding/Executive Officer (CO/XO, n = 5). Further description of the subject demographics can be found in Martin, Sheldon, Kass, Mead, Jones, & Breaux (1998).

Apparatus

The VE testbed was comprised of the following hardware: Dual Processor Octane R 10000 Processors, MXI Graphics, Octane Channel Option, and Indigo^2 Impact R 10000 IDS by *Silicon Graphics, Inc*. Subjects used a VR4 Head Mounted Display (HMD) by *Virtual Research,* and IS600 Inertial Tracker by *Intersence* to view the graphics. The commercial software components were dVise by *Division* and Vega Marine by *Paradigm*. Further specifications can be found in Davidson (1996,1997a, 1997b).

Questionnaires

Subjects were administered six questionnaires: Pre-Questionnaire, Demographics Questionnaire, Pre-Exposure Symptom Checklist, Scenario Review, Post-Exposure Symptom Checklist, and Debrief. The Pre-Questionnaire and Demographics Questionnaire were completed prior to the experimental session. The Pre-Questionnaire solicited comments regarding the critical points of an UNREP, UNREP performance measurements, typical UNREP strategy, and a diagram of the UNREP outlined in the strategy. The Demographics Questionnaire gathered background information on shiphandling, UNREP, and VE experience. The Scenario Review was administered between the performance of the two VE UNREPs to obtain the subject's appraisal of the first UNREP and planned strategy modifications for the second pass. The Debrief was given after the performance of the second UNREP to acquire a comparison of the two UNREPs and usability comments. The results of the usability comments are described in Martin et al. (1998). The Pre- and Post- Exposure Symptom Checklists, an adaptation of the Simulator Sickness Questionnaire (SSQ, Kennedy et al., 1993; Lane & Kennedy, 1988), were used to examine the occurrence of simulator side effects and will be described in a future report.

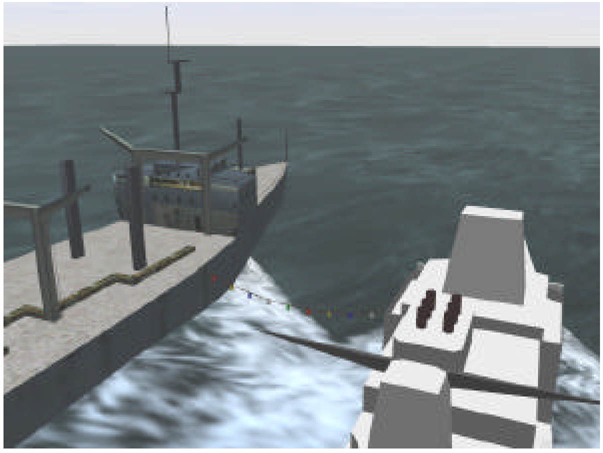
VE UNREP Scenario

The scenario task was to execute an UNREP from the port bridgewing of a guided missile cruiser (CG) and conn the ship alongside a supply ship, maintain the alongside position (at 120 feet lateral separation) for two minutes, and breakaway from the supply ship (see Figure 2 for alongside view). At the scenario start, ownship was positioned 1000 yards directly behind the supply ship, and both ships were traveling on a heading of 130^o at a speed of 15 knots (the UNREP course and speed).

Procedure

Subjects received a review sheet (an informative briefing of the VE ship's characteristics, general reminders regarding hydrodynamic effects, and rules of thumb applicable to UNREP) to study prior to the experiment.

Figure 2. Virtual Underway Replenishment.



The session began with the subject's review of written instructions describing the task and pictures of the location of the supply ship's UNREP station displayed on a PC monitor. The subjects were instructed to issue commands and requests for information as in the real world. These commands and information requests were input to the simulator by an experimenter via keyboard strokes. Replies to commands were made by a pre-recorded speech system, and replies to requests for information were provided verbally by the experimenter. The subjects completed two UNREPs and were given a brief rest period between the UNREPs in which they completed the Scenario Review. It took approximately 1.5 hours to complete the entire experimental session. The first UNREP was considered a practice trial enabling subjects to adapt to the VE. The second UNREP was used for all subsequent analyses.

Following UNREP performance, SMEs were solicited to rate UNREPs presented as plot tracks. Six experienced Surface Warfare Officers rated performance by evaluating a printed track of each subject's UNREP performance. Each track was assigned a rating of 0 to 100. One rater who demonstrated poor internal consistency and poorly correlated with the group was dropped. The mean inter-rater correlation of the remaining five raters = .68; ranging from .56 to .78. The ratings from the five remaining raters were averaged to derive a final performance rating for each UNREP.

Results

The experience level of the sample was diverse, ranging from ensign to commanding officer with a median of 8 years shiphandling experience. The median number of

deployments completed was 4 and the median elapsed time since the last deployment was 3 years. A typical UNREP has an extended duration. Depending on the type of ship, an UNREP can last as long as12 hours (though 1 to 3 hours is more typical), therefore several officers assume the conn during a single evolution. The subject's UNREP experience included completion of a median of 17 approaches, a median of 22 alongsides, and completion of a median of 10 breakaways.
Pursuit of good performance measures began with evaluation of requests for information, engine & rudder commands, & ship dynamic characteristics.

Requests for information (RFI)
Difference comparisons between novice ensigns (no shiphandling experience; n=6) and experienced shiphandlers (n=12) were made for RFI (rudder angle, relative bearing, true bearing, heading, velocity, & range). Novice shiphandlers made significantly more requests for velocity (Novices = 6.3, Experts = 3.0; One-way ANOVA, F=2.40, $p<.05$) and relative bearing (Novices = 11.2, Experts = 3.3; One-way ANOVA, F=6.99, $p<.01$). These differences are consistent with rules of thumb that novices are taught to judge relative positions; experienced shiphandlers rely instead on "seaman's eye" (Crenshaw, 1965) and rarely use these rules and therefore don't make the same RFIs.

In order to determine whether any RFIs were predictive of performance, a linear regression model of SME ratings from RFI was conducted and produced an R=.48 (F=0.59, ns). No individual RFIs were statistically significant in this model. This suggests that RFIs are not effective measures of performance, though they do appear to be indicative of experience.

Engine & Rudder commands
One-way ANOVAs were conducted comparing novice and expert shiphandlers on their cumulative use of shiphandling commands; none of the comparisons on these engines and rudder commands were statistically significant. Furthermore, a linear regression predicting SME ratings from these shiphandling commands was also not significant (R=.47, F=0.91, ns).

Ship dynamics
Candidate measures of ship dynamics as characteristic performance measures were gathered from SME interviews and prior shiphandling dynamics analyses (Martin et al., 1998, Patrey et al., 2000). The most meaningful single relative position, based upon these prior analyses, is within the transitional slide-in phase; in particular, the ship dynamic characteristics (lateral separation, bearing, velocity, & acceleration) at approximately 100 feet astern of the stationkeeping position appears to be the single most distinguishing point. Additionally, measures from the alongside phase for minimum lateral separation (LS), maximum LS, root mean square (RMS) LS, & RMS vertical separation (VS) were included as potentially significant measures.

A linear regression predicting SME ratings from these ship dynamic characteristics was highly significant (R=.98, F=15.76, $p<.001$). In order to create a more parsimonious model, a backward elimination linear regression predicting SME ratings from this host of variables reduced the model to velocity, relative bearing, LS, maximum LS, & RMS LS (R=.92, F=12.99, $p<.001$).

Discussion
Performance measures were successfully identified for virtual UNREP using a traditional approach of identification. Indices of relative position (LS, RMS LS, & maximum LS), relative velocity, and relative bearing significantly predict SME evaluation of performance. Iterative development of the VE coupled with feedback and inputs from SMEs and data analysts enabled the monitoring of salient measures of performance (such as ship dynamics). Furthermore, this has provided a basis for empirically driven performance evaluation.

This clearly demonstrates the functionality of using VE as a tool for deriving performance measures for a real world task. Collecting this quality of data in the real world is a daunting task (though efforts are underway to accomplish this to validate matching between real and virtual UNREPs). While possible to collect this data in the real world, it is difficult and uneconomical to do so, particularly when VE affords an alternative, potentially more effective, method for accomplishing this.

While this particular performance measure derivation effort was primarily driven by a traditional approach to knowledge extraction, virtual data was manually processed with standard statistical methods to glean performance measures that were not wholly apparent from SME interviews. This highlights the need, for at least some types of task, such as those heavily perceptual in nature and not easily verbalized, for additional methods of knowledge elicitation.

Data and cognitive model driven approaches were discussed as potential methods of facilitating and streamlining the knowledge acquisitions process. Currently, both approaches are being investigated for virtual UNREP. Fuzzy logic, as a data driven approach, and COGNET (Cognitive Network of Tasks, Chi Systems Inc.), as a cognitive modeling approach, are the platforms of choice for virtual UNREP and will provide some guidance as to the value in using these powerful tools for performance measure extraction.

This is likely where one of VE's great potential can be realized — as effectual and inexpensive generators of performance indicators, monitors of performance, and ultimately providers of performance evaluation. As these data mining cognitive modeling tools continue to develop, their integration within VE, particularly VE training systems, may prove to be the cornerstone in the revolution in training.

References

Cowden, A.C., Burns, J. J., & Patrey, J. (2000). *Data-driven knowledge engineering.* To be presented at the Interservice/Industry Training, Simulation, & Education Conference.

Crenshaw, R.S., Jr., CAPT, U.S. Navy (Retired). (1975). *Naval Shiphandling*. Annapolis, Maryland: Naval Institute Press.

Davidson, S. (1996). *Software Design of a Virtual Environment Training Technology Testbed and Virtual Electronic Systems Trainer* (Technical Report 96-002). Orlando, FL: Naval Air Warfare Center Training Systems Division.

Davidson, S. (1997a). *Design of an Open Water Shiphandling Software Testbed* (Technical Report 97-007). Orlando, FL: Naval Air Warfare Center Training Systems Division.

Davidson, S. (1997b). Development of a virtual environment software testbed using commercial off the shelf software components. In Proceedings to NATO VR Conference, (December, 1997), Orlando, FL.

Kennedy, R.S., Lane, N.E., Burbaum, K.S., & Lilienthal, M.G. (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3 (3),* 203-220.

Lane, N.E. & Kennedy, R.S. (1988). *A New Method for Quantifying Simulator Sickness: Development and Application of the Simulator Sickness Questionnaire (SSQ).* (Technical Report EOTR 88-7). Orlando, FL: Essex Corporation.

Martin, M.K., Sheldon, E., Kass, S., Mead, A., Jones, S., & Breaux, R. (1998). Using a virtual environment to elicit shiphandling knowledge. *Proceedings to '98 I/ITSEC Conference*, Orlando, FL, December 1998.

Patrey, J., Sheldon, E. M., Breaux, R.B., & Mead, A. M. (2000). Quantifying performance of a dynamic shiphandling perceptual-action task. Technical Report in preparation.

Pew, R. W., Mavor, A. S., *Modeling human and organizational behavior: application to military simulations*, National Academy Press, Washington, D.C. 1998.

This page has been deliberately left blank

Page intentionnellement blanche

Appropriate Use of Virtual Environments to Minimise Motion Sickness

Willem Bles & Alexander H. Wertheim[1]
TNO Human Factors
Kampweg 5
3769 DE Soesterberg
The Netherlands

1. Introducton

With the current fast rate of technological developments and the high requirements for training with sophisticated apparatus, the military has become more and more involved in working with simulators. The term "simulator" here means: a systems that has the potential to create sensations of passive or active self movement in a simulated environment. This definition of the term "simulator" not only applies to the traditional flight simulators, both with and without a moving base, but also to Virtual Environments (VE) set-ups implemented in Head Mounted Display (HMD) systems, which no doubt will become part of future flight training programs.

Apart from the obvious usefulness of such simulators, they also have a serious disadvantage: it turns out that they expose users to discomforting and unwanted side-effects, that might well affect training efficiency. One of the most important and well known problems is that these simulators often induce motion sickness, which severely interferes with behaviour and thus with training. Motion sickness causes lowering of motivation, usually resulting in a considerable slowing down of work rate, a disruption of continuous work, or even its complete abandonment. In fact, motion sickness in simulators is currently the main factor limiting the use of simulators.

There are various kinds of motion sickness, such as air sickness, sea sickness, car sickness, space sickness, and some people may even get sick in trains or elevators.
Simulator sickness is basically a form of motion sickness. It has been defined as motion sickness which occurs in a simulator, but which would not occur in the real world in the same circumstances as those which are simulated [28]. For instance, if a person gets sick in an aeroplane and also in a simulator, which validly mimics the flight movements, then this would not classify as simulator sickness. We only speak of simulator sickness if that person would become sick in the simulator but not in the aeroplane. The same reasoning applies to motion sickness in virtual environments.

In order to be able to minimise the incidence of motion sickness in virtual environments, it is necessary to understand the reasons for simulator sickness, and thus for motion sickness in general. Therefore we will briefly review our present view on motion sickness. This will then allow us to understand why some factors are important to lower the motion sickness incidence in virtual environment applications.
Finally we will discuss other, often related, human factor problems that happen frequently in virtual environments, such as headaches, eye strain and after-effects, and mention what might be done to minimise these effects.

2. Motion sickness in general

Motion sickness may vary among subjects: within individuals, there is no direct correlation between sensitivity to various forms of motion sickness. Sensitivity to any particular form of motion sickness also varies largely among humans. Moreover, motion sickness may develop fast or slow. Women are generally somewhat more sensitive than men. There seems to be an effect of age as well. Sensitivity for motion sickness is very low with children a few years old, then increases and at old age decreases again [36].
It is known that, after its initial rise, motion sickness eventually decreases with time despite ongoing motion exposure. This adaptation may take a few hours up to a few days, as with sea or space sickness. But again, the time it takes for the symptoms to disappear differs among individuals. With approximately 5% of humankind adaptation does not take place at all.

All this makes it difficult to understand the nature of the provocative motion stimulus. In a series of experiments, carried out in a Ship Motion Simulator (SMS), McCauley et al. suggested that it is mainly the vertical component of ship motion that causes sea sickness [34]. For sinusoidal vertical motion they found motion sickness to be most prominent between 0.05 to 0.8 Hz (maximum at 0.2 Hz) and with amplitudes of over 1 m/s^2, the incidence of motion sickness increasing further at higher amplitudes. On the basis of their data these authors developed a descriptive mathematical model of sea sickness [31]. More recently another mathematical motion sickness incidence model has been proposed by Griffin, allowing also for complex vertical motion patterns [23] (for comparison of these two models, see [16, 17]). These models became the basis for the international standards. The main premise of these descriptive models is that varying vertical accelerations are an important factor in the generation of motion sickness.

[1] For contact with authors: bles@tm.tno.nl and wertheim@tm.tno.nl

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

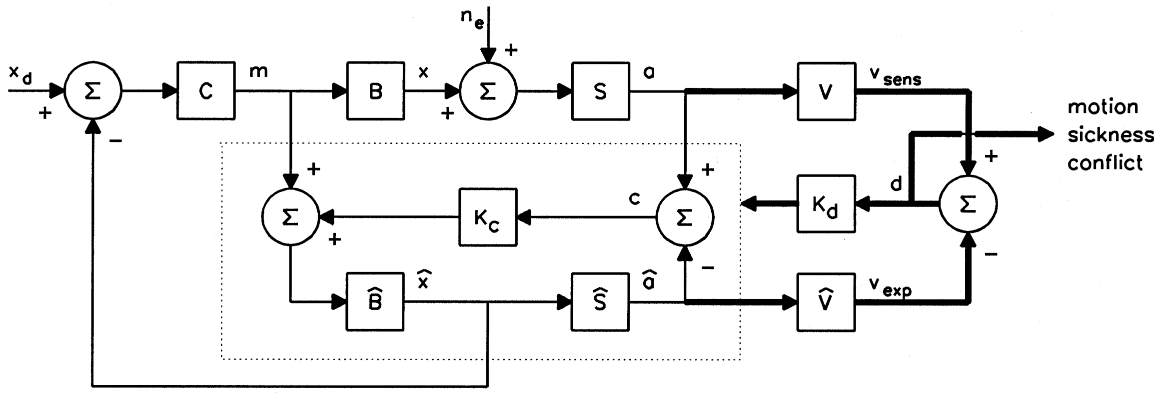


Fig. 1 Instead of the conflict vector **c** from the Oman model [35], the subjective vertical motion sickness model considers the vector **d** to be the conflict vector for generating motion sickness. The modules V are necessary for the computation of the subjective vertical (thick lines). The dotted lines represent the internal model.

It has also been shown that motion sickness may develop as a result of horizontal linear movements [22, 25].
Furthermore, the notion that only vertical movements are sea sickness provoking, has been challenged in a series of experiments by Wertheim with a ship motion simulator [43]. He observed motion sickness even when vertical movements — which, according to the above mentioned mathematical theories, were too weak to generate motion sickness — were accompanied by low frequency pitch and roll motions. Head movements still further increased the motion sickness incidence [42]. Ergonomic measures to counter motion sickness at sea included the design of a working place such that head movements could be minimised [4].
Head movements which changed the orientation of the head with respect to gravity also proved to be very provocative in subjects who had been submitted previously to constant hyper gravity in a human centrifuge (2–3 g for 1.5 hrs [7, 33]).

These examples illustrate the view that the vestibular system plays a crucial role in the generation of motion sickness [21, 36]. In fact, it has long been known that the one necessary requirement for any kind of motion sickness is a functioning vestibular apparatus. People who do not have a functioning vestibular apparatus (because of particular illnesses) simply cannot become motion sick [e.g. 29].
However, vestibular-visual interactions are also very important in provoking or preventing motion sickness: the driver of a car does not get sick, whereas the passenger reading in the back seat may have a fair chance of getting sick on a curved road. Somatosensory-vestibular interactions also prove to be important in the incidence of motion sickness as was demonstrated with (Pseudo-)Coriolis effects [6]. Especially with VE these interactions are very important.

This is not the proper place to present a detailed description of how the vestibular apparatus works. Many good texts on the subject ore available elsewhere (e.g. Guedry [24], or Howard [26]. Here it suffices to note that the central role of the vestibular system is recognised in what are currently the most well known explanatory theories of motion sickness, like the theory of intersensory mismatch [35, 36].
A more specific version of this theory assumes that motion sickness results from only one mismatch, the one between the expected vertical and the vertical as determined on the basis of the incoming sensory information [8]. There are some other alternative theories based on ecological perspectives [39, 44], and there are ideas about cognitive influences on motion sickness [19], but here we will focus primarily on the view that motion sickness arises when there is a mismatch in the determination of the gravity representation.

According to the sensory mismatch theory from Reason and Brand [36], motion sickness occurs when the sensory systems provide the brain with more than one kind of self-motion information which do not match each other. This could be either an intra- or an inter-sensory conflict.
The sensory mismatch theory offers some remedies for motion sickness. For example, in the case of a ship at sea, the incidence and severity of sea sickness under deck should be reduced when the visual system is provided with an optic pattern which remains stable not relative to the eyes, but relative to the real world. This was proposed by Bittner & Guignard [4] and it fits the experience that standing on deck with view of the horizon is less provocative than standing under the deck. In fact there have been some attempts to investigate possible motion sickness reducing effects of an artificial horizon [10, 38].

3. The subjective vertical mismatch concept

Although many examples of conflicts between and within sensory systems can be described, leading to disorientation and motion illusions indeed, there is plenty of evidence that motion sickness is primarily provoked in those situations where the determination of the subjective vertical, the internal representation of

gravity, is challenged. Therefore the sensory rearrangement theory on motion sickness was redefined to: "All situations which provoke motion sickness are characterised by a condition in which the sensed vertical as determined on the basis of integrated information from the eyes, the vestibular system and the non-vestibular proprioceptors is at variance with the subjective vertical as predicted on the basis of previous experience" [8, 9]. In Fig. 1 this concept is illustrated. Since with this model in principle motion sickness incidence can be described for every stimulus condition, such an approach would be more useful than the descriptive models as discussed above for sea sickness, since these descriptive functions only apply to particular stimuli which have to be determined first. We therefore explain this model in more detail for the situation of walking towards a certain position.

In Fig. 1 we see that, in order to obtain the desired position $\mathbf{x_d}$, muscle activity (m) is generated leading to a position $\mathbf{x}$ due to the body dynamics (B). This signal, together with the external noise n_e, is detected by the senses (S) resulting in sensory information $\mathbf{a}$. The internal model consists of the same components (indicated with a hat) and computes the expected sensory information $\hat{\mathbf{a}}$. Differences between the vectors $\mathbf{a}$ and $\hat{\mathbf{a}}$ are calculated, and are fed back into the system. In this way an optimal estimate of the actual walking path can be obtained.

The Subjective Vertical conflict model extends the Oman model [35] with a network V which constructs the sensed vertical, $\mathbf{v_{sens}}$, based on the incoming sensory information. Similarly, in the internal model a network $\hat{V}$ is added which constructs the expected vertical, $\hat{\mathbf{v}}$ or $\mathbf{v_{exp}}$, based on previous experience and expectation. The difference vector $\mathbf{d}$ between $\mathbf{v_{sens}}$ and $\mathbf{v_{exp}}$ is used to update $\mathbf{v_{exp}}$, and is in our view the conflict vector which generates motion sickness [8] (see Fig. 1).

For analysis of the provocativeness of motion conditions it is of great importance to know how the representation of the vertical is accomplished [8, 11].

This is in fact the basic vestibular problem for the central nervous system. In Fig. 2 it is shown how this could be accomplished on the basis of psycho-physiological evidence. The vestibular (semi-circular canals, SCC, and otoliths, OTO), the visual (VIS) and the somatosensory (SOM) system all provide information on spatial orientation. In order to obtain only one unique spatial orientation it is assumed that all this sensory information is integrated (INT) into basically three signals, indicating the sensed rotation (SR), the sensed translation (ST) and the sensed vertical (SV) as shown in Fig. 2 [8].

The integration of rotatory motion information is rather straightforward, because the sensory systems provide complementary information. A more complex problem for the central vestibular system is to extract the gravity information out of the sensed gravito-inertial force vector. In view of normal human movements and locomotion, it was hypothesised that low-pass filtering (LP) of the signal representing the gravito-inertial force vector could preserve gravity. This is a sensible approach, provided that the angular motion information helps to compensate for the consequences of fast head tilts. Mathematically this compensation is accomplished by a transformation R of the co-ordinate frame with the otolith vectors, over the angle of the head tilt indicated by the rotation sensors. Such a manipulation keeps the input to LP unchanged, the sensed vertical after the head tilt being determined by the rotatory motion information due to the inverse transformation R^{-1} as shown in Fig. 2.

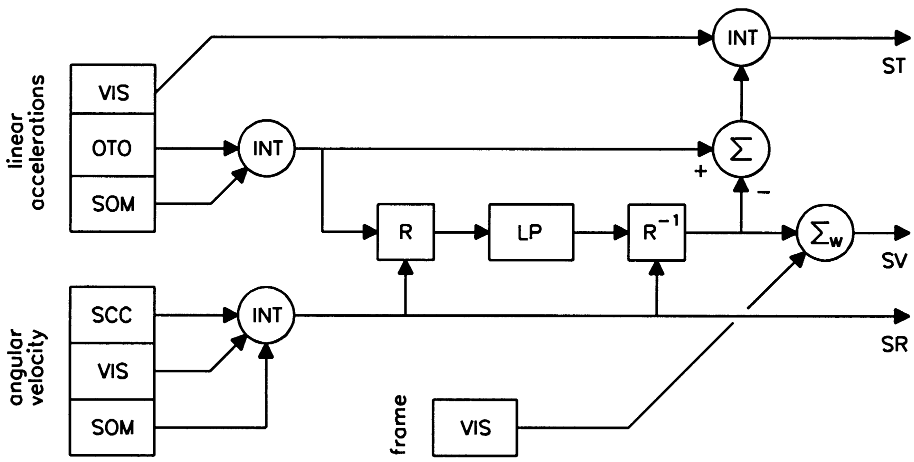


Fig. 2 Integration of sensory information.

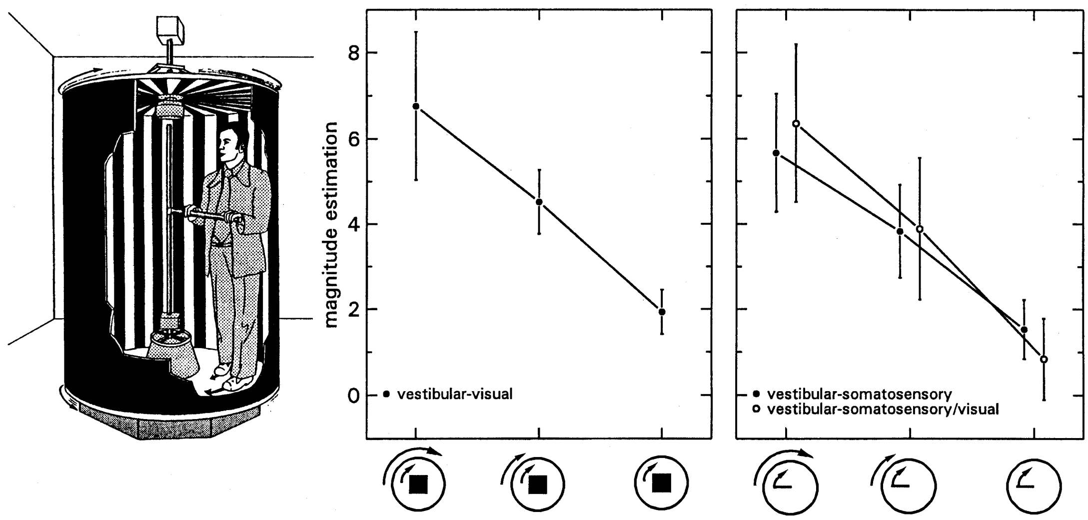


Fig. 3 Differential effects of congruent and incongruent visual and somatosensory motion stimulation on the magnitude of the vestibular Coriolis effect (5 is the standard magnitude of the discomfort of the vestibular Coriolis effect).

It is assumed that the internal model uses a similar neural network as on the sensory side, the values of the different parameters being determined by previous experience. To illustrate this point, the observation is of interest that experienced pilots are suffering less from motion sickness in real flight than student pilots, whereas they are more prone to simulator sickness than student pilots. The internal model of an experienced pilot apparently has parameter settings that match quite well the motion signals which are sensed by the sensors during real flight, but they do not match to the information as sensed by the sensors in, for instance, a fixed-base simulator environment. For student pilots the argument goes the other way around: they have no particular experience as for the in-flight environment. Thus, in the simulator the match is better then during real flight, where they sense motion signals which are not expected.

To summarise, difference vectors between sensed and expected linear and rotatory motion are not a trigger for motion sickness: this may only result in disorientation. Only differences between the sensed and expected vertical provoke motion sickness.
This is illustrated in modern architecture where fully listed buildings are popping up more and more: In a stationary listed environment (visual frame information not coinciding with the gravity vector) head movements were found to be provocative to motion sickness. This was described by Kitahara & Uno [30] and we confirmed this observation: Walking in a stationary listed environment (max. 20 degrees) made about 10% of the subjects motion sick within 15 minutes. Especially the turning around proved to be provocative [11]. In this condition it is noteworthy that a stationary subject doesn't get motion sick, despite the continuing conflicting information from the visual frame and the otoliths about the direction of the vertical. According to the SV conflict model, the sensed and the expected attitude converge due to the feedback in this situation: Only when the subject starts to move around, differences are to be expected between these two vectors. This is a common observation in many motion sickness provoking surroundings: Moving around or making head movements enhances motion sickness (see section 2).

4. Factors causing nausea in virtual environment simulators

Somatosensory-visual-vestibular interactions. With the principle of the subjective vertical mismatch one can analyse the different virtual environments concepts on provocativeness for motion sickness. To illustrate the meaning of this concept, the results of laboratory experiments which are of direct relevance for the use of HMD and VE system concepts, are shown in Fig. 3. The results stem form experiments done by Brandt et al. [14] and Bles [5].
In these experiments the magnitude of the nausea of the Coriolis effect obtained by lateral head tilt during constant velocity rotation at 60 □/s was studied under different visual and somatosensory stimulus conditions. The pure vestibular Coriolis effect, head tilt in darkness, served as a reference and had a magnitude of 5. It shows that the Coriolis effect is minimal if there is sight on the earth-stationary visual surround. This is comparable to walking conditions with a HMD with a perfect earth-stationary virtual environment. The nausea increases if the visual surround rotates together with the chair, which is compatible to rotating with a HMD with a head fixed display. If the surround rotates with twice the chair velocity, the nauseating effect of a head tilt is very strong. This demonstrates what happens if the HMD provides non-earth referenced motion information. Inspection of the right frame in Fig. 3 indicates that

manipulation of somatosensory motion information as obtained by stepping in circles in darkness provides similar results as manipulation of the visual information when sitting. If in these conditions the visual and optokinetic motion information is combined, the modulation of the vestibular Coriolis effect is even more pronounced (Fig 3 right frame, open dots). This shows how important it is to take into account the somatosensory information, if present, otherwise the analysis may lead to predictions which are completely different from the experimental data. The Subjective Vertical model as shown before fully accounts for these experimental results [6]. The SV model also perfectly applies to the concept of the closed cockpit aircraft [11].

Fixed base vs. moving base. In order to minimise simulator sickness for HMDs with virtual environment applications, the same rules apply as for fixed and moving base simulators. Moving in a virtual environment of a HMD may be accomplished by turning or walking on a treadmill, or by means of a joy-stick. These changes of propagation means, together with irregular motion velocity patterns using the joy-stick may even be more demanding from the human equilibrium system than a normal 6DoF flight simulator. In fact, keeping in mind the frequency characteristics of the different parts of the equilibrium system, the model in Fig 1 may help to analyse the stimulus patterns on their provocativeness to motion sickness. It is no surprise that a HMD training facility on board of a moving platform with motion which has absolutely nothing to do with the training scenario, is due to be more provocative than on a non-moving platform.

Destabilisation of the visual world. If one makes a head movement while wearing a HMD, the image in front of the eyes will move with the head. In other words, in such situations the visual world loses its stability [40].
An additional complication here is that when we make a rotatory head movement, the eyes rotate in the head in counter direction. This so called Vestibulo-Ocular Reflex (VOR) normally serves to maintain ocular fixation on an object in our environment during head movements. The VOR is very fast and has a latency of approximately 10 ms. However, to maintain ocular fixation when the object moves with the head (as in a HMD) the VOR must be suppressed. The necessary enervation of the ocular musculature is relatively slow and frequency specific. With head movements up to 1Hz the VOR can be properly suppressed, but at higher frequencies the VOR dominates, blurring the visual image on the retinas and causing visual discomfort. If the blur stems from very fast retinal motion its direction cannot be perceived, which may have consequences for the computation process as indicated in Fig. 2.

Image magnification (or minimisation). Similar problems may occur when an outside image, projected inside an HMD (e.g. the image of a night vision goggle), is magnified or minimised. Normally the VVOR (VOR with full sight on the visual surround) has a gain of 1, which means that the velocity and amplitude of this reflexive eye movement is equal to that of the counterdirective head movement. If the head movement is fed back to move a magnified image in a HMD in counter direction to the head, the velocity of the image shift is higher than expected, while with a minimised image it is lower. This means that the visual information contributing to the computation in Fig. 2 may not properly match the vestibular inputs in the computations, which may also lead to discrepancies in the determination of the representation of gravity. Such situations resemble the case where one scans the scenery with binoculars in which case the visual image moves across the retinas with a much higher speed than is normally the case during head movements. The same happens when wearing new spectacle glasses. But since glasses are usually worn continuously the visual vestibular interaction may adapt back to normal in due time. However, as long as such adaptation is not complete nausea might persist. Unless one wears a HMD for quite a long time similar adaptation may not easily be obtained. Thus it is recommended not to use a magnification or minimisation factor in the design of VE or HMD visuals with outside image representations.

There is another discomforting problem related to image magnification or minimisation. The point is that when we stand upright we normally make small body movements (body sway). Here the visual system helps. It feeds these small retinal image shifts back to the system which maintains body posture. When those image shifts do not really correspond to how the body really moves (because of their optical magnification or minimisation), they are still fed back to our musculature with which we maintain our postural equilibrium. Thus we may end up making much larger body sway motions, which poses a threat to our postural stability and may create feelings of insecurity with respect to our equilibrium (in fact it is this mechanism which causes fear of heights — in which case the image movements have become disproportionally small, because of the very far distance of objects in the visual environment [13].

Time delays. In many VE simulations head movements are fed back to the visual display, with the purpose of moving the image across the display in the direction counter to the head movement. This should ensure that the virtual environment remains stationary relative to earth (i.e. relative to gravitation and compass-fixed) during head movements. However, in many simulators, including HMD-systems, this coupling is less than perfect, which may cause severe nausea. The point is that the visual image must move across the display surface in precise temporal synchrony with the movements of the head. Otherwise a phase difference between the visual and vestibular inputs to the CV and SV occurs which may cause them to deviate from each other, causing severe nausea. However, it always takes

time to record (and filter) head movements and to calculate the movements of the image inside the HMD on the basis of these records. This manifests itself in a temporal delay of the required visual image changes, especially with very large and detailed displays. During the delay period there is a large discrepancy between visual and vestibular information: the head movements are properly registered by the vestibular system, but the visual world moves *with* in stead of *against* the head. Even with delays as brief as 46 ms, the resulting visual-vestibular mismatch, which may easily cause a CV vs. SV mismatch, may already be extremely nauseating [20, 27].

This reasoning is in line with empirical results from recent experiments in which the gain and phase relations of visual and vestibular information were manipulated separately, using an artificial environment set up mounted on a sled for linear motion [32]. The data clearly suggested that phase differences are much more provocative than gain differences, and that, in contradistinction to visual phase-*leads* (relative to the vestibular stimulus), small visual phase-*lags* are already highly provocative.

Vection. Visually induced sensations of self-movement, known technically as "vection", are of course key phenomena in simulators. However, since visual suggestions of self-motion may easily affect the Sv through the integration INT with the SCC and SOM information (see Fig. 2), they always form a potential risk of motion sickness. In this section we will review the properties of visual displays and images that affect vection, and which thus have to be considered in evaluating the risk for the development of nausea in simulators.

Screen size. Vection is strongest with peripherally moving visual flow fields. Hence, large screens carry higher risks of motion sickness. With full-field flow fields almost everyone will experience strong sensations of vection. Thus as a general rule, the smaller the visual image (or display) the lower the chance of motion sickness. From laboratory experiments it has been concluded that the risk of vection is minimal with images extending a visual angle less than approximately 30°. A normal standard 17 inch computer screen viewed at a distance of 50 cm encompasses 34° and therefore will not easily generate vection.

Foreground/background. A necessary condition for vection to occur is that the inducing visual pattern is perceived and interpreted as a background. Normally, when walking past an object which we fixate with our eyes, its background moves in our visual periphery, the central area of the visual field is occupied by the retinally stationary object. However, when we move in a vehicle (e.g. a car), the situation is reversed. Here the peripheral parts of our visual field are occupied with objects that remain stationary on the retinas (e.g. the hood of the car, the frame of the windshield, the dashboard etc). In such situations vection is caused by image motion across the central area of the visual field. Experiments have shown that such centrally evoked vection is possible only if the visual flow is perceived as background, that is, as further away in depth than the stationary objects in the periphery. Hence, in exception to the above mentioned rule, visual patterns covering small visual angles may still evoke vection if they are perceived as a background. Thus small displays in simulators, which simulate "out-of-the-window" views may facilitate vection.

Pattern motion. As should be clear by now, moving visual patterns always carry with them a certain chance that vection develops. With a constant velocity pattern vection normally develops with a latency between up to 20 seconds (depending on various stimulus parameters) after which vection velocity does not increase any further and the pattern appears earth stationary. At this point vection is said to be saturated. The forcefulness with which vection is experienced and the perceived velocity of vection depend not only on the size of the vection inducing pattern, or on whether or not it is perceived as a background, but also on its velocity. Perceived vection velocity increases with the velocity of the stimulus pattern up to approximately 60º/s, after which it is reduced rapidly and the visual pattern is perceived as unstable or just moving.

Vection also depends on the motion frequency of the inducing pattern. As mentioned above, its latency can be relatively long, implying that low frequencies are more powerful than high frequencies. With sinusoidal pattern motion frequencies up to 0.1 Hz vection can normally be induced. At higher frequencies vection rapidly decreases. Thus if one wants to prevent vection it is important to keep this cut-off frequency of 0.1 Hz in mind.

5. Other discomfort factors in head mounted displays

Image flicker. Typical computer work complaints such as eye-strain, visual fatigue, headache and blurred vision, are common also when working with HMDs. The reason for these complaints are not always clear, but one of the causes often suggested is image flicker. Our sensitivity for image flicker is higher in the visual periphery than in the central visual field. Causes for image flicker are long times needed for computing the motion of images in the HMD (update frequency), especially when these computations must be carried out on the basis of on-line head movement registrations, and the refresh rate of the particular screen used in the HMD. It is advisable to avoid screens that have a refresh rate of less than 80 Hz. Traditional video screens are too slow (50 Hz). To reduce the risk of perceiving flicker it is also advisable to reduce the luminance of the images in the HMD to less than 50 cd/cm2 and to keep luminance contrasts relatively low as well.

Image acuity and depth perception. Bad image acuity may also yield complaints of headache and eye-strain, especially when text has to be read. Image screens should have a resolution at least comparable with that of a 1024 × 768 pixel 17 inch computer monitor. Traditional video screen technology has too low a resolution to be acceptable in HMDs, especially with wide angle screens.

With 3D VR systems, the two eyes receive separate and slightly different images, which are fused by the brain to perceive depth. It is advisable to facilitate the fusion process as well as possible, by positioning the image optically at 2 to 4 m distance from the eyes. The necessary ocular accommodation is then 0.5 to 0.25 dioptres and the necessary convergence of the eyes then covers 0.9 to 0.4 degrees of visual angle. If the two images are not placed at the correct position relative to the eyes eye-strain will result from the additional oculo-muscular effort required.

To keep a reasonable visual acuity in such 3D VR and HMD systems, the following criteria should apply with respect to corresponding details in the two images (correct adjustments of the rims of the images is less critical):

- The (rotational) difference between corresponding details should not exceed 1°.
- The vertical position of corresponding details should not exceed 0.5°.
- Divergence between corresponding details should be no more that 0.5°.
- The size of corresponding details should not differ by more than 3%.
- The difference in required accommodation of the two eyes should not exceed 0.25 dioptres.

Smoothness of image motion. To avoid headaches and eye strain in simulators it is necessary that smooth visual motion will indeed be perceived as smooth. This is not always the case. The same factors apply here as those which cause flicker. When calculations necessary for generating moving images take relatively much time (low update rate), or screen refresh rate is low, the movements will be seen as consisting of small steps. This is visually quite discomforting.

Motion parallax. On the flat surface of visual displays there is no real depth. It must be simulated. Not only by proper perspectives which change during simulated ego motion, but more importantly, by concurrent relative motion between the objects in the surroundings (motion parallax). If motion parallax is not properly programmed, it may create impressions of self motion which do not properly fit vestibular cues from the motion base. For example, most simulator systems use visual display systems in which the movements of the vehicle (e.g. an air plane) are fed back to change the visual image on the display in such a manner that it appears stationary with respect to the real world. However, if the head movements of the individual inside the simulator are not fed back to affect the image on the display in a similar manner, the concurrent vestibular sensations may not always match the changes in the image. For example, imagine a person inside such a simulator who moves the head closer to a visual display unit that is supposed to simulate a window through which a visual outside scene is seen. The eyes then get closer to the screens. In normal situations more of the visual environment will then become visible from behind the rims of the window and the size of the retinal images of far away objects will not change much. Conversely, if such a forward head movement is made in a simulator, where the observer's head position is not fed back to the image on the screen, no new parts of the environment will become visible from "behind" the rims of the screen and the images of all virtual objects will be enlarged equally on the retinas, whatever their distance. Therefore the changes in the visual information will not match the vestibularly sensed head movements. This may cause visual discomfort and, if lasting long enough, eye-strain and headache. If that visual-vestibular mismatch includes aspects of the subjective or sensed vertical, a risk of motion sickness may evolve as well.

Control device system lag. When using a computer mouse, a joy stick, roller ball or any other control device to affect the image on a visual display in a simulator which is used in an interactive man-in-the-loop mode, performance may be affected when delays between the action and its effect on the screen become too large. Such delays are not discomforting in the sense that they might cause motion sickness, headaches etc, but they may well have a deteriorating effect on tracking and steering performance.

No hard limits can be given for maximum lags because they also depend on the kind of vehicle model used in the simulator (see for a review: Ricard [37]). However in general steering performance is assumed to deteriorate when control device system lags increase beyond 100 ms [1, 2], while lags over 300 ms may induce oscillations [3]. With respect to normal computer use, lag times for the use of a mouse should not become larger than 50 ms, while the lag between pressing a key on a key-board and the appearance of a letter on the display should not be longer than 100 ms (DERA defence standards [18]).

After-effects. When trainees spend many hours inside a simulator there is a risk of after- effects once they exit the simulator. Such after-effects include not only a continuation of nausea, but also postural imbalance and headaches (see for a review: Wertheim [41]). They may have negative effects on performance in normal everyday behaviour (e.g. driving), or may aversely affect special skills such as are involved in flying an air plane. This issue has been recognised in the literature as having juridical consequences for those responsible for simulators and trainees. They might find themselves liable if trainees cause accidents after a simulator training. Only recently has research started on such after-

effects and currently there is not much specific information available as to their exact nature and the risks involved. However, after-effects may last for many hours [3].

6. Conclusions

Head Mounted Displays still easily provoke discomfort. The known visual problems in using HMDs which are due to the technical limitations of the display and computing limitations, will most probably be solved by technical improvements in the near future. As long as that is not the case, the factors described in section 5 should be taken into account.

In developing HMD application concepts one should be aware of the motion sickness consequences of orientation cues which lead to false visual verticals, because of the fact that a discrepancy between the sensed and expected representation of gravity is considered to be the primary motion sickness provoking conflict. Qualitative analysis with the model on the provocativeness of the application taking into account what is known on the sensory interactions is very useful already. Quantitative analyses by Bos & Bles [12] have shown that the model accounts for the sea sickness data of O'Hanlon and McCauley [34]. This is a very promising accomplishment, since the international standards (see section 2) are based on descriptive models.

References

1 Allen, R.W. & DiMarco, R.J. (1984). Effects of transport delays on manual control system performance. Paper presented at 20th Annual Conference on Manual Control, NASA Ames Research Center, 12-14 June, 1984. System Technology, Inc., 13766 South Hawthorne Boulevard, Hawthorne, CA., USA.

2 Bailey, R.E., Knotts, L.H., Horowitz, S.J. & Malone, H.L. (1987). Effect of time delay on manual flight control and flying qualities during in-flight and ground-based simulation. AIAA paper 87-2370 at the AIAA Flight Simulation Technologies Conference, New York. American Institute of Aeronautics and Astronautics. Calspan Advanced Technology Center, Buffalo, New York, USA.

3 Baltzley, D.R., Kennedy, R.S., Birbaum, K.S., Lilienthal, M.G. & Gower, D.W. (1989). The time course of postflight simulator sickness symptoms. Aviation Space and Environmental Medicine, 60, 1043-1048

4 Bittner, A.C. & Guignard, J.C. (1985). Human factors engineering principles for minimizing adverse ship motion effects: theory and practice. Naval Engineers Journal 97-4:205-213

5 Bles, W. (1981). Stepping around: circular vection and Coriolis effects. In: Long, J.; Baddeley, A., eds. Attention and Performance IX. Hillsdale, NJ: Lawrence Erlbaum Ass.

6 Bles, W. (1998). Coriolis effects and motion sickness modelling. Brain Research Bulletin, 47 (5): 543-549.

7 Bles, W., Bos, J.E., Furrer, R., De Graaf, B., Hosman, R.J.A.W., Kortschot, H.W., Krol, J.R., Kuipers, A., Marcus, J.T., Messerschmid, E., W.J.Ockels, W.J.Oosterveld, J.Smit, A.H.Wertheim & C.J.E.Wientjes (1989). Space Adaptation Syndrome induced by a long duration +3Gx centrifuge run. Rept. IZF-1989-25, TNO Human Factors Research Institute, Soesterberg, The Neterlands.

8 Bles, W., Bos, J.E., De Graaf, B., Groen, E.L. & Wertheim, A.H. (1998). Motion sickness: Only one provocative conflict? Brain Research Bulletin, 47(5):481-488.

9 Bles, W., Bos, J.E. & Kruit, H. (2000). Motion Sickness. Current Opinion in Neurology 2000, 13, p. 19-25.

10 Bles, W., De Graaf, B., Keuning, J.A., Ooms, J., De Vries, J. & Wientjes, C.J.E. (1991). Experiments on motion sickness aboard the M.V. "Zeefakkel". Rept IZF-1991-A-34, TNO Human Factors Research Institute, Soesterberg, The Netherlands.

11 Bles, W. & Tielemans, W.C.M. (2000). Motion sickness consequences of flying closed cockpit aircraft. Countering the directed energy threat: are closed cockpits the ultimate answer? RTO Meeting Proceedings 30 (RTO-MP-30 AC/323(HFM) TP/10). 22:1-6

12 Bos, J.E. & Bles, W. (1998). Modelling Motion Sickness. RTO/HFM's Aircrew safety Assessment 26:1-6

13 Brandt, Th. (1999). Vertigo: its multisensory syndromes. -2nd ed. Springer Verlag, Berlin Heidelberg New York

14 Brandt, Th., Wist, E. & Dichgans, J. (1971). Optisch induzierte Pseudocoriolis-Effekten und Circularvektion. Archiv Psychiatrische Nervenkrankheiten, 214, 365-389

15 Conklin, J.E. (1957). Effect of control lag on performance in a tracking task. Journal of Experimental Psychology 53(4):261-268.

16 Colwell, J.L. (1989). Human factors in the naval environment: a review of motion sickness and biodynamic problems. DREA Technical memorandum 89/220, Canadian National Defence Research Establishment Atlantic. Dartmouth.

17 Colwell, J.L (1994). Motion sickness habituation in the naval environment. DREA Technical Memorandum 94/211, Canadian National Defence Research Establishment Atlantic. Dartmouth.

18 DERA Defence Standard 1996. Defence standard 00-25 (part 13)/Issue 1, 24 May 1996; Human Factors For Designers Of equipment, Part 13: Human Computer Interaction. DERA, Malvern, Worcs, WR14-3PS, UK.

19 Dobie, T.G. & May, J.G. (1994). Cognitive-behavioral management of motion sickness.

Aviation Space and Environmental Medicine, 65, 10, section II, C1-C20.
20 Draper, M., Viire, E., Furness, T.A. & Parker, D.E. (1998). Theorized relationship between vestibulo-ocular adaptation and simulator sickness in virtual environments. Proceedings of the International Workshop on Motion Sickness: Medical and Human Factors, Marbella, Spain.
21 Eyeson-Annan, M., Peterken, C., Brown, B. & Atchison, D. (1996). Visual and vestibular components of motion sickness. Aviation, Space and Environmental Medicine, 67, 10, 955-962.
22 Golding J.F. & Kerguelen, M. (1992). A comparison of the nauseogenic potential of low frequency vertical versus horizontal linear oscillation. Aviation, Space and Environmental Medicine, June issue, 491-497
23 Griffin, M.J. (1990). Handbook of human vibration. Academic Press, London.
24 Guedry, F.E. (1974). Psychophysics of vestibular sensation. In: H.H. Kornhuber (ed): Handbook of sensory physiology Vol 6/2, Springer, NY. 3-154.
25 Horii, A., Takeda, N., Morita, M., Kubo, T. &nd Matsunaga, T. (1993). Acta Otolaryngologica, Suppl 501, 31-33
26 Howard, I.P. (1986). The vestibular system. In: Boff, K.R., Kaufman, L. and Thomas J.P. (eds): Handbook of Perception and Human Performance, Vol I: Sensory processes and perception. John Wiley, NY.
27 Kalawsky, R.S. (1993). The science of virtual reality and virtual environments. Addison-Wesley Publishers.
28 Kennedy, R.S., Berbaum, K.S., Allgood, G.O., Lane, N.E., Lilienthal, M.G. & Baltzley, D.R. (1988). Ethiological significance of equipment features and pilot history in simulator sickness. NATO-AGARD conference proceedings No. 433: Motion cues in flight simulation and simulator induced sickness. Neuilly sur Seine, France, 1-1/1-22
29 Kennedy, R.S., Graybiel, A., McDonough, R.C. and Beckwith, F.D. (1968). Symptomology under storm conditions in the North Atlantic in control subjects and in persons with bilateral labyrinthine defects. Acta Oto-laryngologica, 66, 533-540
30 Kitahara, M. & Uno, R. (1967). Equilibrium and vertigo in a tilting environment. Annals Otol (St Louis) 76:166-178.
31 McCauley, M.E., Royal, J.W., Wylie, C.D., O'Hanlon, J.F. & Mackie, R.R. (1976). Motion sickness incidence: exploratory studies of habituation, pitch and roll, and the refinement of a mathematical model. Human Factors Research Inc. Technical Report 1733-2.
32 Mesland, B.S. (1998). About Self-motion Perception. PhD thesis, University of Utrecht. The Netherlands.
33 Ockels, W.J.R., Furrer, R. & Messerschmid, E. (1990). Space sickness on earth. Experimental Brain Research, 79, 61-663.
34 O'Hanlon, J.F. & McCauley, M.E. (1974). Motion sickness incidence as a function of the frequency and acceleration of vertical sinusoidal motion. Aerospace medicine, 45, 366-369.
35 Oman, C.M. (1982). A heuristic mathematical model for the dynamics of sensory conflict and motion sickness. Acta Oto-Laryngologica, suppl. 392
36 Reason, J.T. & Brand, J.J. (1975). Motion Sickness. Academic Press. London.
37 Ricard, G.L. (1994). Manual control with delays: A Bibliography. Computer Graphics, 28, 2, may 1994, 149-154.
38 Rolnick, A. & Bles, W. (1989). Performance and well-being under tilting conditions: the effects of visual reference and artificial horizon. Aviation Space and Environmental Medicine, 60, 779-785.
39 Stoffregen, T.A. & Riccio, G.E. (1986). Out of control: an ecological perspective on motion sickness. Paper presented at the Fall meeting of the International Society for Ecological Psychology, October 18, Philadelphia, PA, USA.
40 Wertheim, A.H. (1994). Motion perception during self motion: the direct versus inferential controversy revisited. Behavioral and Brain Sciences 17: 293-355.
41 Wertheim A.H. (1999). The assessment of aftereffects of real and simulated self-motion: motion sickness and other symptoms. Rept. TNO-TM 1999-A074, TNO Human Factors Research Institute, Soesterberg, The Netherlands.
42 Wertheim, A.H., Heus, R., & Vrijkotte, T.G.M. (1995). Human energy expenditure, task performance and sea sickness during simulated ship movements. TNO- Report TM-1995-C-29; TNO Human Factors Research Institute, Soesterberg, The Netherlands.
43 Wertheim, A.H., Wientjes, C.J.E., Bles, W. & Bos, J.E. (1995). Motion sickness studies in the TNO-TM Ship Motion Simulator (SMS). TNO-Report TM-1995 A-57, TNO Human Factors Research Institute, Soesterberg, The Netherlands.
44 Yardley, L. (1992). Motion sickness and perception: a reappraisal of the sensory conflict approach. British Journal of Psychology, 83, 449-471.

This page has been deliberately left blank

Page intentionnellement blanche

Keynote Address 2: Available Virtual Reality Techniques Now and in the Near Future (Unclassified and for distribution to all NATO nations)[1]

Grigore C. Burdea
Rutgers University
Human-Machine Interface Laboratory/CAIP
96 Frelinghuysen Rd.
Piscataway, NJ 08854-8088
USA

Summary

This paper presents available virtual reality technology as well as technology that is projected to become available to NATO in the near future. Areas discussed are new PC technology (graphics rendering and wearable computers), personal and large-volume displays, large volume tracking, force feedback interfaces, and software toolkits. PCs presently render millions of polygons/sec. Their reduced cost makes possible the distribution of virtual environments at many sites and in many countries. Large-volume displays are more expensive, but allow more natural user interactions. They do require large-volume tracking that is fast and accurate. Haptic interfaces are a recent class of input/output devices that increase simulation realism by adding the sense of touch. This comes at a cost of more computing power and better physical modeling. The modeling and programming needs of virtual reality are met by software toolkits designed for such simulations.

1. Introduction

Virtual reality technology has experienced significant advances in the late nineties, and now has many characteristics that may be exploited by the military. Virtual reality has the potential to significantly reduce training costs and the risk to him. It also has the potential to reduce *team* training costs, allowing multi-national organizations, such as NATO, to have a unified training system, without a unique training location. Virtual reality, as a computerized training environment, allows transparent gathering of data, and the remote access to such data, at a much smaller time interval, and resolution than allowed by manual data collection methods. For all these reasons it is important to inform the military decision-makers of what technology and methods are available today, or what will become available in the near future.

This report is based on the keynote address given by the author at the NATO Workshop that took place in April 2000 in Hague. Then, as now, the time and space available for such a review are limited. When trying to condense all this material, which can easily take a Semester to teach in college, certain things had to be omitted. Thus the present review does not cover networked communication as it applies to shared VR, nor does it cover human factor trials of VR technology. Such topics are covered in companion papers. Emphasis here is on commercial off-the shelf technology, or technology that is close to commercialization. Many deserving research projects are omitted here, as a matter of practicality. The interested reader who wants more information on such research should consult the open literature, such as the *Proceedings of the IEEE Virtual Reality Conference* series (formerly VRAIS), and other such publications.

Section 2 of this report presents significant changes in the computing platforms that are (or may be) used in VR. Section 3 describes the displays that output the graphics scene to the user, whether such displays are personal or large-volume. Large-volume displays, in turn, require large-volume trackers, which are the subject of section 4. Section 5 presents the newer haptic interfaces, which bring more realism to the simulation by allowing the user to touch and feel virtual objects. The modeling libraries needed by modern VR simulations (including haptics) are detailed in section 6. Section 7 concludes this report.

2. The PC Revolution

Probably one of the most important changes that has influenced the VR arena in recent years is the tremendous increase in PC-based graphics rendering speed. The closing gap between inexpensive PC-based graphics and the high-end SGI engines is clearly illustrated by Figure 1.

The measure of performance used for comparison here is the number of polygons rendered by the computer in unit time. When dividing this number by the scene complexity, one obtains the screen refresh rate in frames/second (how many snapshots of the virtual scene the computer can render per unit time). The more complex the scene, the less frames/second, which in turn can result in a disturbing saccadic graphics [Burdea & Coiffet, 1994].

[1] Based on the author's presentation at RTA/HFM Workshop 007, The Hague, Netherlands, 13-15 April.

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

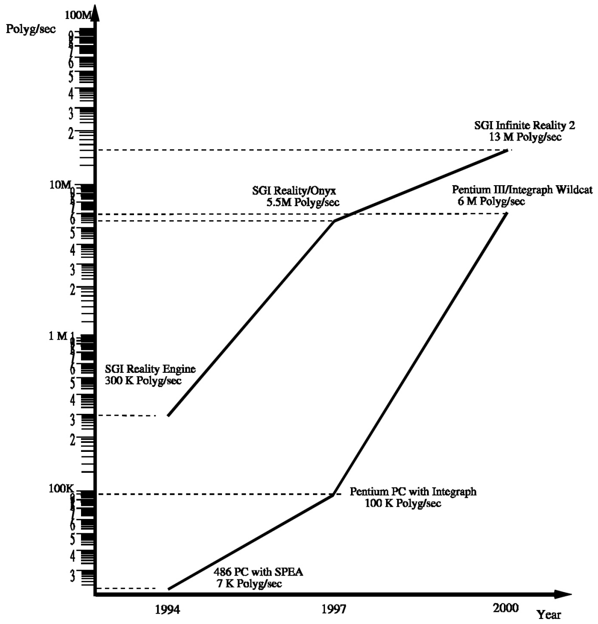


Figure 1: SGI graphics vs. PC-based graphics

In 1994 a 486 processor PC with SPEA FIRE board was capable of 7,000 polygons/sec. A modern Pentium III PC with Wildcat graphics board can do 6,000,000 polygons/ sec, and costs only 6,000 dollars or so. During the same time the performance of high-end graphics workstations produced by SGI rose from 300,000 polygons/sec. on a Reality Engine in 1994 to 13,000,000 polygons/sec. today on a multi-pipe Infinite Reality 2 [Real Time Graphics, 2000]. While its performance is twice that of the fastest PC rendering board, its price is two to three hundred thousand dollars, which makes it affordable to only a few! By significantly improving performance, while actually reducing costs in the late nineties, the PC industry made possible the much-desired widespread use of desktop 3-D graphics.

The second important change in the computer industry is the tendency to miniaturize the computer, to the point that it becomes wearable on the user. Figure 2 shows just such an example, namely the Mobile Assistant IV® produced by Xybernaut Co. (Fairfax VA, USA). It consists of a CPU unit with a Pentium processor and simplified keyboard, a head-mounted display, a microphone for voice input, and a camera worn on the user's head. By coupling this with wireless communication, the user gets freedom of motion within the range of the wireless transmitter, and as a function of battery life.

User freedom of motion is very important to the VR application designer, because it increases the naturalness of the interaction, and thus the feeling of immersion that the user has. At the present time the Mobile Assistant does not have sufficient computing power to incorporate graphics real-time rendering. Such a capability is expected to appear in subsequent models of the device.

Figure 2: Mobile Assistant IV® wearable computer. Courtesy of CAIP Center, Rutgers University. Reprinted by permission

3. Graphic Displays

Another important component of VR systems are the graphics displays, which present the computer, rendered scene to the user. Such displays may be classified as personal displays, for a single user, and large-volume displays, which allow several users to view the same scene in a given location. Both types of displays have advanced significantly in the past decade, as will be described next.

3.1 Personal displays

The most prevalent type of personal display available in the nineties were head-mounted displays (HMDs), which projected the image close to the user's head. Early HMDs were very bulky and heavy, weighing over two kilograms in the case of the VPL "Eyephone." Their resolution was poor (360×240 pixels) owing to the LCD technology of the time. Compared to this, modern HMDs, such as the SONY Glasstron® shown in Figure 3, have an SVGA resolution (832×624 pixels). The improvement in image resolution was coupled with a dramatic reduction in weight (120 grams for the Glasstron). Unfortunately, the necessary miniaturization means that the user's field of view (FOV) is small (30×22 degrees) compared to the Eyephone FOV of 90×60 degrees. Recently SONY has announced it will stop producing Glasstrons. Its logical replacement is the Olympus Eye-Trek HMD (37×22 degrees) weighing a little over 100 grams [Olympus, 2000].

Figure 3: The SONY Glasstron Courtesy of InterSense Co. Reprinted by permission

The user’s natural field of view is 180 degrees horizontal and almost as much vertical. The human vision system, unlike the HMDs, has an uneven resolution over its FOV. The highest resolution is in a central “foveating area,” while the retina has much lower resolution away from the foveating area. By rendering the image at constant resolution the computer essentially wastes pixels, since the eye cannot see them. Eye trackers allow computers to detect where the user focuses on an image. It is then possible to render the corresponding virtual scene in high resolution, and the rest of the scene in lower resolution. A review on the state-of-the-art in eye tracking can be found in [Isdale, 2000]. Figure 4 shows an HMD retrofitted with an eye tracker.

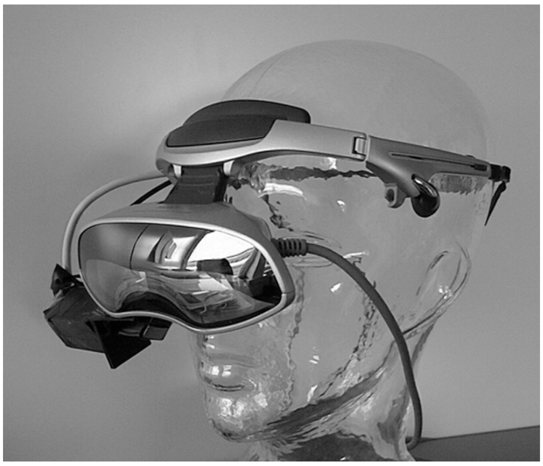


Figure 4: The SONY Glasstron fitted with an eye tracker. Courtesy of *VR News.* Reprinted by permission

Military reconnaissance training applications can benefit from a “customized” HMD, such as the V8 Binoculars (Virtual Research Systems Inc., Santa Clara CA, USA) shown in Figure 5. These binoculars integrate dual LCD displays, with VGA resolution, and a FOV of up to 60 degrees. Its optics allows individual focus adjustment, and its weight is 680 grams. By integrating a position tracker (discussed later in this report), the computer senses the 3-D aim of the binoculars and displays the corresponding scene in real time.

Figure 5: The V8 Binoculars HMD. Courtesy of Virtual Research Systems Inc. Reprinted by permission

Other types of graphics displays, available today, are “virtual windows” and auto-stereoscopic displays. The WindowVR® produced by Virtual Research Systems Inc., is shown in Figure 6. In has a flat-panel display (a touch-sensitive display in some versions) with handles and suspension cable. A tracker inside the display allows the computer to change the scene and give the user the sensation of looking at a virtual world through a window. Buttons on the handles allow actions and navigation within the VR simulation.

Figure 6: The WindowVR®. Courtesy of Virtual Research Systems Inc. Reprinted by permission

Auto-stereoscopic workstations, such as the ones produced by Dimension Technologies Inc. (Rochester NY, USA), use backlighting of a flat panel to produce a stereo image. As seen in Figure 7, the image appears to float in space, without the need for HMDs. Its resolution is 1280×1024, which is superior to that of LCD-based

displays [Dimension Technologies Inc., 2000]. Unfortunately, the stereo image can be seen from only a small viewing volume and the brightness of the image suffers owing to the lighting scheme used. Thus graphics appears dim when compared HMDs or active glasses (discussed later in this report).

Figure 7: An auto-stereoscopic workstation. Courtesy of DTI Inc. Reprinted by permission

3.2 Large-volume displays

Large-volume displays offer a much larger stereo viewing area, high resolution, and a way for many participants to view and interact with the same virtual scene. One class of large-volume displays is "virtual workbenches," such as the one shown in Figure 8. It uses a CRT projector and mirrors to "place" the stereo scene on top of its table. The integration of its projector within the display table makes for a compact design, and the tilting mechanism can change the user's viewing cone. The Baron can tilt from fully horizontal to fully vertical, which transforms it essentially in a "virtual wall" type display. Future designs will replace the CRT technology with much brighter digital mirror technology. Then it will be possible to use such displays without having to reduce the room ambient lighting level.

Figure 8: The BARCO Baron® 3-D display. Courtesy of BARCO Co. Reprinted by permission

Figure 9 shows a marine amphibious landing exercise scene produced by a workbench-type display [Hix et al., 1999]. The usual 2-D military symbols were replaced by 3-D icons of trucks, airplanes, ships, etc., shown on a 3-D terrain map. Such a scene is much easier to comprehend, and may reduce errors in a high stress combat situation. Furthermore, the use of 3-D icons coupled with haptics (not used in this particular training scenario) opens the way for a different kind of C&C interaction.

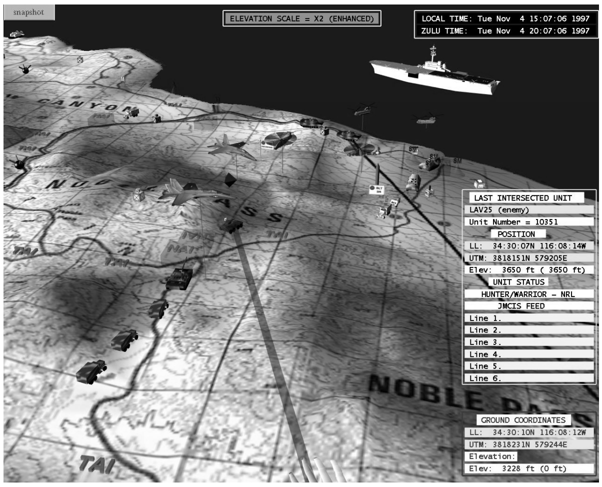


Figure 9: Sea Dragon Marine landing exercise. Courtesy of the Naval Research Laboratory, Washington DC. Reprinted by permission

Using a haptic glove (discussed later in this report) the military commander may then be able to grasp and feel such 3-D objects. The force feedback addition to the simulation has at least two important advantages for the military decision-maker. First, he knows he has complete and unique control over the unit whose symbol he grasped. This is true even if he momentarily looks away from the screen. Second, the hardness of the symbol can give him valuable information on the unit's state of readiness/strength level. A tank 3-D icon that feels soft may indicate that unit is at half strength, due to losses. A tanker plane that feels hard may indicate that it is full of fuel, etc.

An example of a C&C application using a haptic glove is the system demonstrated by the CAIP Center at Rutgers University, and shown in Figure 10 [Medl et al. 1998]. It consists of a distributed architecture, with a multi-modal interface. The user gives voice commands that are detected by a microphone array placed on top of a PC. He can select and move military symbols on a map using either an eye tracker, or a force feedback glove (Rutgers Master glove [Burdea, 1996]). The New Jersey National Guard, with little prior training, tested the system successfully in 1997.

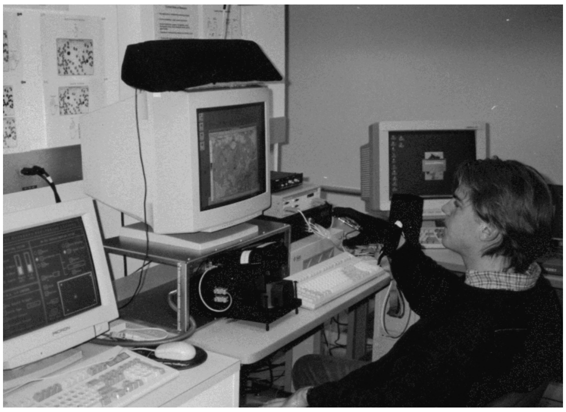


Figure 10: Multi-modal interface C&C exercise. Courtesy of the CAIP Center, Rutgers University. Reprinted by permission

A larger type of display than the workbench is the CAVE® stereo display made by Fakespace Systems (Ontario, Canada). As shown in Figure 11, the CAVE consists of multiple wall-type displays assembled in a cube geometry. Each wall has its own CRT projector, driven by a separate graphics pipe of a multi-processor high-end SGI or equivalent computer. The user enters the CAVE and is looking at the display walls through "active" stereo glasses, such as those shown in Figure 12. Infrared emitters located in the corners of the CAVE control the opening and closing of shutters incorporated in the stereo glasses. They alternately block the view of each eye, which allows the brain to register the two images rendered by the computer separately and create the stereo effect.

With his FOV filled by the graphics the CAVE user feels immersed in the virtual world. Furthermore, the work volume in which the user sees stereo and can interact with virtual "floating" objects is much larger than for a workbench. These advantages come at a price, as the cost of the CAVE is five times that of a workbench display. T this is added the cost of the high-end graphics computer, bringing the system close to one million dollars at the time of this writing.

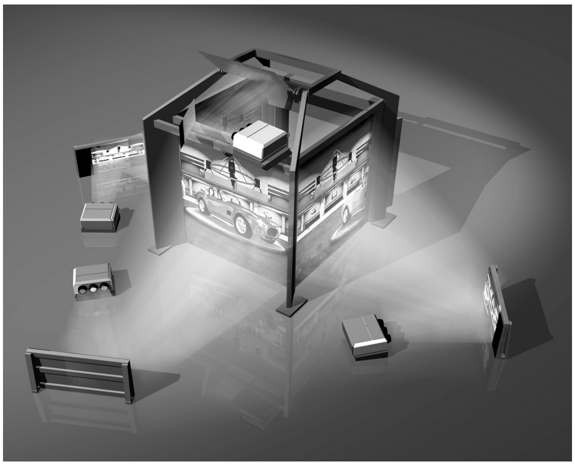


Figure 11: The CAVE stereo display. Courtesy of Fakespace Systems Inc. Reprinted by permission

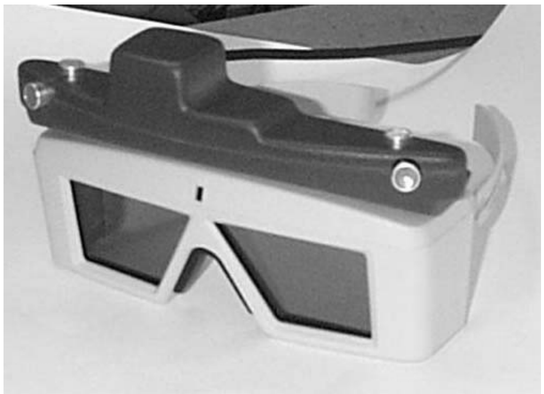


Figure 12: Stereo "active" glasses fitted with the InterSense tracker. Courtesy of InterSense Co. Reprinted by permission

Recently Fakespace Systems introduced the "Re-configurable Advanced Visualization Environment" (RAVE) shown in Figure 13. Unlike the CAVE, which has a fixed geometry, RAVE can change its configuration depending on the user's needs. Thus its 3 m × 2.9 m × 3.7 m modules can be assembled to form a straight wall geometry, where three display units are side-to-side. Other available configurations include a u-shape, or a cube (CAVE-type geometry). Alternately, it can separate itself into two half-cube independent displays. As expected, the cost of RAVE surpasses that of the CAVE.

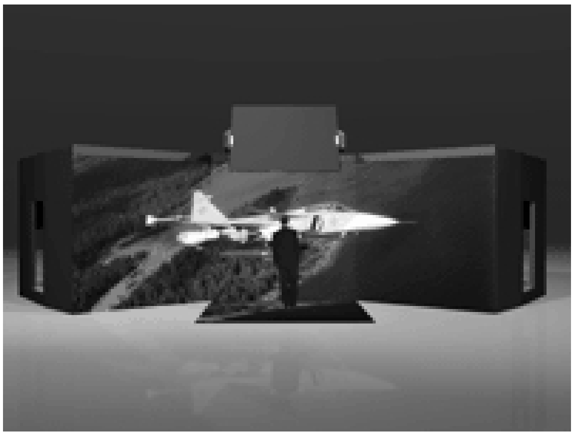


Figure 13: The RAVE re-configurable stereo display. Courtesy of Fakespace Systems Inc. Reprinted by permission

4. Large-Volume Tracking

The user's ability to see graphics that fill most of his FOV is a good start towards a more immersive virtual environment. Another important requirement is to allow the user to interact with virtual objects he sees. Thus the computer needs to know as accurately as possible the current 3-D position of the user's hand(s), head, or whole body within this large working volume.

4.1 Magnetic tracking errors

Computers determine the user's position by interpreting data fed by 3-D trackers worn on the body. The overwhelming majority of today's trackers are

electromagnetic ones, consisting of a stationary source of pulsating magnetic fields, one to several receivers (coils) worn by the user, and an electronic control box. The voltages induced in the receivers are transformed in absolute position/orientation values by the control box, and then sent to the computer running the simulation.

An example of high-end magnetic tracker is the MotionStar® wireless tracking suit produced by Ascension Technology Co. (Burlington VT, USA), shown in Figure 14. The suit incorporates 20 magnetic tracker receivers placed at critical locations on the user's body, such as the wrist, ankle, hip, etc. The receivers are wired and the electronic control/communication box worn on a backpack. Owing to its own power supply (a battery with two-hour life), the suit can work independently and furnish up to 100 readings/sec. within three meters from the tracker source. Such a range would accommodate two RAVE modules, if placed side-by-side, with the source centrally located.

Figure 14: The MotionStar® wireless tracking suit. Courtesy of Ascension Technology Co. Reprinted by permission

There is however a problem with all magnetic trackers, which affects their accuracy. This is due to interference from other magnetic fields, or from metallic objects. Such problems were reported with the MotionStar® [Marcus, 1997], but also with the Polhemus LongRanger® (Colchester VT, USA) [Trefftz & Burdea, 2000]. Figure 15 shows the magnitude of the error vector for a LongRanger® installed on a wooden tripod in the Human-Machine Laboratory at Rutgers University. The tripod allowed the height of the tracker source to be varied, while precise position of the receiver was measured mechanically. The errors grew geometrically with the distance from the tracker source, as expected. However, errors also varied depending on the source height above the floor. The most accurate measurements were obtained when the source was at 1.68 m above the floor. Errors grew when the source was too close to either the ceiling or to the floor, owing to the metallic beams used in the laboratory room construction. Additional experimental measurements showed that the metal in the large-volume display (in this case a BARCO Baron workbench) introduced more tracking errors.

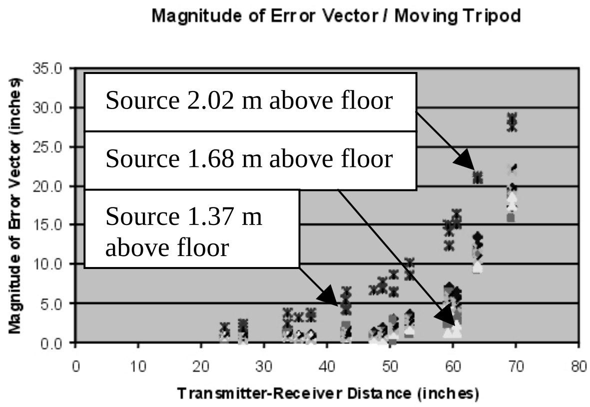


Figure 15: The Polhemus LongRanger tracking errors [Trefftz & Burdea, 2000]

The above findings, and those of others, point out the inadequacy of magnetic trackers when working in typical large-volume display environments. Thus one is left with two alternatives. The first is to build a special structure, designed from the start to house large-volume displays and the related trackers, and to redesign the display to reduce the amount of metal. The second, and an easier alternative, is to change the tracker.

4.2 Inertial/ultrasonic trackers

In recent years a new generation of trackers has become commercially available. These are hybrid 3-D position trackers, such as the IS-600 shown in Figure 16, manufactured by InterSense Inc. (Burlington MA, USA). They use a combination of inertial and ultrasonic sensing technology, with the inertial component used for position measurements, and the ultrasonic component used to provide a zero position and to correct for drift. One or more inertial cubes are placed on the user, or on his interface, together with sonic disks (as previously shown in Figure 12 for active glasses). The inertial cube signal is read by an electronic box, which also drives ultrasonic receivers placed on the ceiling in a cross configuration. Since these trackers do not use magnetic fields, they are immune to the type of interference associated with magnetic trackers.

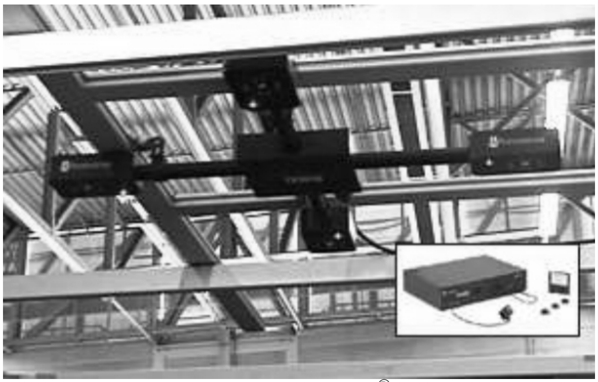


Figure 16: The InterSense IS-600® inertial/ultrasonic tracker. Courtesy of InterSense Co. Reprinted by permission

A recent addition to the InterSense tracking family is the IS-900 LAT (large-area-tracker) [InterSense, 2000]. It can extend its 6 m × 6 m × 3 m standard tracking volume to a maximum tracking area of 900 m^2 using up to 24 expansion hubs. Its measurement accuracy, resolution and latency are better than for magnetic trackers.

5. Haptic Interfaces

Another important change taking place in current VR technology is the addition of haptic feedback, namely tactile and force feedback. Tactile feedback gives the user the ability to touch and feel the smoothness of virtual object surfaces, their temperature, slippage, and contact surface geometry. Force feedback conveys information on object weight, inertia, mechanical compliance, degree of mobility, viscosity, etc. The addition of haptic feedback clearly increases simulation realism in general. Furthermore, haptic feedback allows object manipulation in occluded, foggy or dark virtual environments, a task that would otherwise be difficult or even impossible to complete.

5.1 General-purpose haptic interfaces

Haptic interfaces may be classified as general-purpose ones, which can be used for many tasks (including military ones), and special-purpose haptic interfaces, which are designed specifically for military applications.
An example of a general-purpose force feedback interface is the PHANToM® arm Desktop produced by SensAble Technologies Co. (Woburn MA, USA), and shown in Figure 17. The interface measures the position and orientation of the stylus 1000 times/sec, and applies forces of up to 10 N to the user's hand in response to actions in the virtual environment. The high bandwidth of the PHANToM allows it to combine force with tactile feedback, such that the roughness or stickiness of a surface can be simulated as well.
A typical application developed for the PHANToM is "digital sculpting," as illustrated in Figure 17. The user is presented with a block of "digital clay," which he deforms, sculpts, polishes, using the stylus. The user feels the resistance of the material, as well as the influence of the change in virtual tool to which the stylus is mapped.

Figure 17: The PHANToM® desktop force feedback arm. Courtesy of SensAble Co. Reprinted by permission

Once the 3-D model is sculpted, its files can be downloaded to a NC mill or similar equipment, to build an actual prototype. This is also applicable to the weapon design cycle, speeding up its mock-up phase.
Another use of the PHANToM is in mine detection training, an application being currently developed by the French Ministry of Defense (see companion paper by Todeschini). The force feedback arm integrated with this system is designed to replicate the tactile sensation the trainee uses to detect a mine. Since in actual operations such a task must have a 100% rate of success, it is clear that a realistic trainer should be useful. The difficulty in realizing such a system is to realistically replicate the dynamic force "signature" associated with various mines and ground conditions.

Figure 18: Digital sculpting with force feedback. Courtesy of SensAble Co. Reprinted by permission

One drawback of the PHANToM arm is that it is not able to provide finger-specific forces, such as those present in dexterous tasks, when contact is at the fingertip. Such tasks could be assembly training, servicing of military hardware, or training in explosive handling. For such instances a better haptic interface is a

force feedback glove, such as the CyberGrasp® glove produced by Virtual Technologies Inc. (Palo Alto CA, USA), shown in Figure 19.

Figure 19: The CyberGrasp glove in a CyberPack configuration. Courtesy of Virtual Technologies Co. Reprinted by permission

The glove consists of a CyberGlove [Kramer et al., 1991] used for position measurements on which is retrofitted a force feedback exoskeleton driven by cables. The tendons are routed to an electronic control box housing electrical actuators and communication hardware. The force output is about 16 N per finger, which is larger than the PHANToM output. Unlike the PHANToM, which sits on a desk, and limits freedom of motion, the CyberGrasp glove is worn. Furthermore, the CyberPack® configuration places the control box in a backpack, such that the user can walk around and grasp objects and feel their hardness. Its limiting factors then are weight, (which can lead to user fatigue) and the range of the tracker measuring wrist 3-D position.

Another limitation of the CyberGrasp haptic glove is the lack of force feedback to the wrist. Thus grasped objects seem weightless, with no inertia and no mechanical restraints. Recently Virtual Technology announced the CyberForce® haptic interface shown in Figure 20. It consists of a six degrees-of-freedom force feedback arm connected to the back palm. By combining wrist force feedback with the force feedback glove, the ability to simulate weight and inertia are added while the user preserves his hand dexterity [Kramer, 2000]. Furthermore, there is no need for a wrist position tracker, since the force feedback arm measures wrist position faster and without metallic interference. Unfortunately, the dimensions of the arm limit the user's freedom of motion. Furthermore, the overall system control becomes much more complex, which may lead to system instabilities.

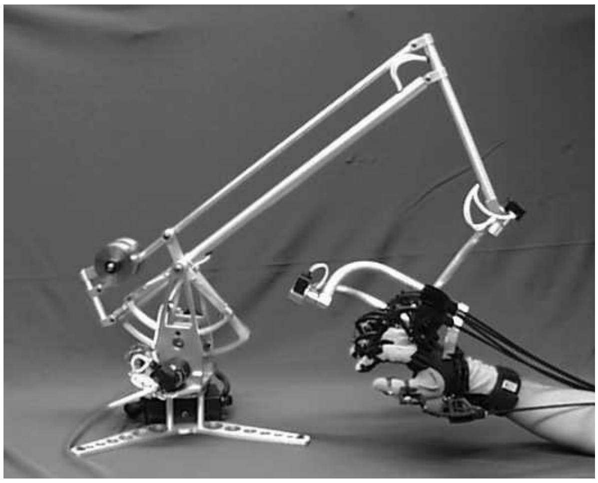


Figure 20: The CyberGrasp glove in a CyberForce configuration. Courtesy of Virtual Technologies Co. Reprinted by permission

In certain military applications of VR, such as infantry training, there is a need to simulate running, or walking uphill, or through uneven terrain. In such cases haptic feedback to the body becomes important in order to have realistic training. One system that addresses these needs has been recently developed by Sarcos Co (Salt Lake UT, USA) and the University of Utah [Hollerbach et al., 1999]. As shown in Figure 21, the user is located in front of a three-wall display filling most of his FOV and stands on a treadmill. By tracking his walking/running on the treadmill, the computer updates the virtual scene accordingly. A force feedback arm is attached to the user's torso through a harness. The arm applies resistive and inertial forces to simulate uneven terrain and other effects. A rope attached to the ceiling prevents injury in case of tripping and falling.

Figure 21: The treadport VR system. Courtesy of University of Utah CS Dept. Reprinted by permission

Recently, Japanese researchers proposed the replacement of the treadmill approach with an "active floor", as shown in Figure 22 [Noma et al., 2000]. The floor is composed of modular actuator tiles that can change slope under computer control. The user's motion is tracked by

a vision system, and the tiles actuated as needed to replicate uneven terrain. Thus, unlike the walking-in-place paradigm of treadmill systems, the active floor approach allows natural walking over the whole surface of the floor. There is no need for a force feedback arm attached to the user's back, and no need for a safety rope. The limitation in this case is the size and amount of slope that can be produced by the active tiles.

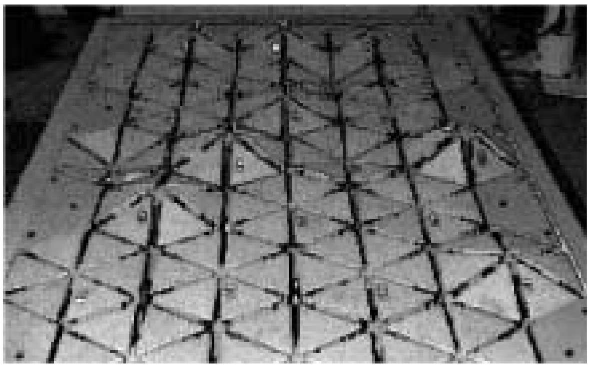


Figure 22: The active floor VR system [Noma et al. 2000]. © IEEE. Reprinted by permission

5.2 Special-purpose haptic interfaces

All the haptic interfaces presented so far are general-purpose, since they can be used in military applications but were not specifically designed for such. By contrast, special-purpose haptic interfaces are designed from the start to provide force/touch feedback to military VR tasks. An example is the Stinger trainer prototype developed at TNO (The Hague, The Netherlands) [Jense, 1993], shown in Figure 23. It consists of a plastic mock-up of the missile launcher, which is instrumented to track the user's aim, and to sense when switches are depressed. Furthermore, a virtual environment showing the enemy aircraft is presented to the trainee on an HMD. The advantage of this system is that a much more compact set-up replaces the classical large-dome training system. Furthermore, all user actions are stored transparently and his performance data is available on the computer. The force feedback sensation is produced naturally by the plastic mock-up, without need for more expensive (and heavier) hardware. The system is now being used in training the German Air Force, as described in the companion paper by Reichert.

Another example of special-purpose haptics is the anti-tank missile trainer system recently developed by the Fifth Dimension Technologies Co. (Pretoria, South Africa), which is shown in Figure 24. It uses a mock-up of the rocket launcher, similar to the TNO Stinger trainer, which provides direct tactile feedback. Other similarities include the used of a HMD to display the virtual battlefield to the trainee, and a 3-D tracker to determine his direction of view.

Figure 23: The Stinger VR training prototype Courtesy of TNO, The Netherlands. Reprinted by permission

Figure 24: The anti-tank VR training prototype Courtesy of 5DT Co., Pretoria, South Africa. Reprinted by permission

Another type of special-purpose haptic interface is the parachute-training simulator developed by Systems Technology Inc. (Hawthorne CA, USA). As shown in Figure 25, the system uses a full-size parachute harness, and an HMD showing a detailed 3-D jump scene (insert). The scene moves in response to either head motion, or the toggle of the parachute harness [Systems Technology Inc. 2000]. Wind effects are added, to train the jumper in coping with adverse landing conditions. Playback of user actions and instructor actions are used to help acquire the necessary skills.

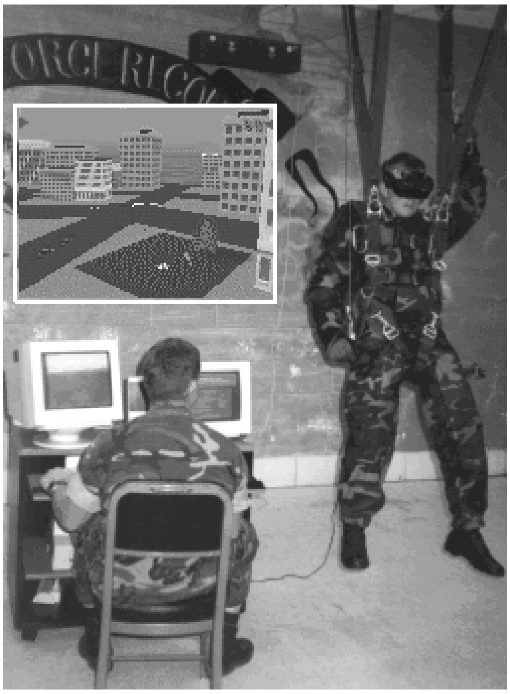


Figure 25: The VR parachute training system. Courtesy of Systems Technology Inc. Reprinted by permission

6. Modeling Tools

So far this report has reviewed the computing hardware and the interfaces available to develop VR applications. The third element needed is a VR toolkit, i.e. software libraries specifically developed for programming virtual environments. Such toolkits offer certain advantages to the developer, namely drivers for most VR I/O devices, certain 3-D graphics routines, ease of portability, etc. In turn VR toolkits can be classified as general-purpose and special-purpose libraries.

6.1 General-purpose Modeling Tools

The most used VR programming toolkit today, by far, is “WorldToolKit” (WTK), produced by Sense8, a division of Engineering Animation Inc. (Ames IA, USA). It consists of over 1000 C/C++ object-oriented functions, which are executed, in an infinite loop during the simulation. An example of a scene created with WTK is the tank interior simulation shown in Figure 26. By importing CAD files, doing smooth shaded graphics, textured surfaces, dynamic effects, WTK allows very realistic simulations to be created.

Another facility provided by WTK (in its “World-up” version) is graphics programming, as shown in Figure 27. Thus the kinematics dependencies and other virtual object characteristics can be easily specified using a scene graph. At run time the software goes through the nodes of this scene graph.

For all its advantages WTK has at least two disadvantages, namely cost and short-lived releases. The license cost for WTK is an order of magnitude more than for widespread PC software, reflecting the small market for VR products. This is aggravated by numerous releases, which many times are not compatible with earlier ones. As such a military application developed with an earlier release may not run when the library is updated (currently WTK is at release 9).

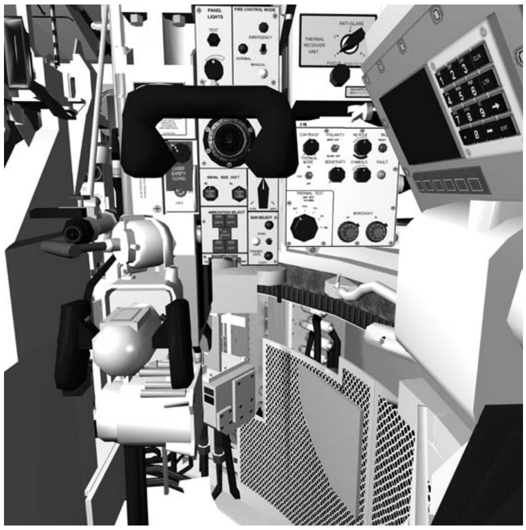


Figure 26: The tank interior created with WTK. Courtesy of EAI Co. Reprinted by permission

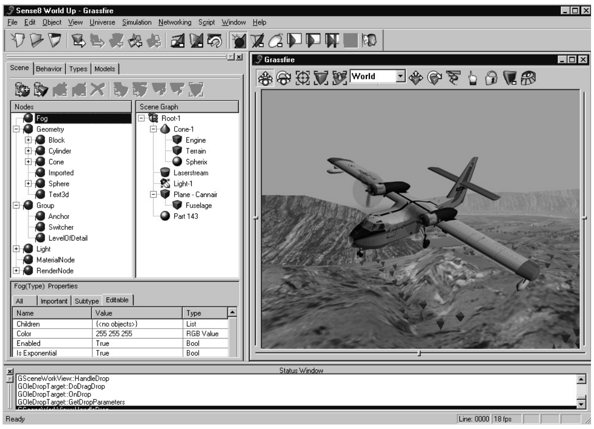
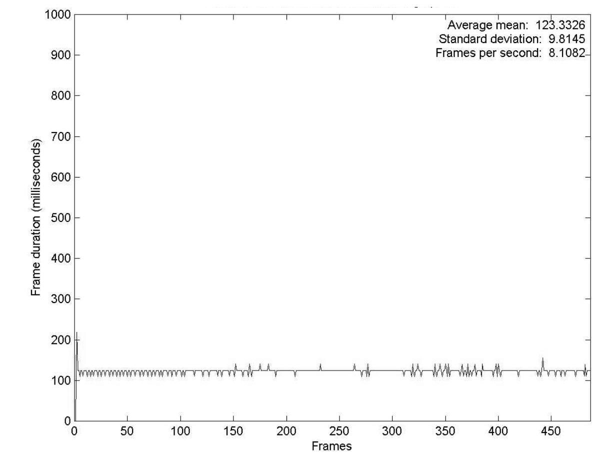
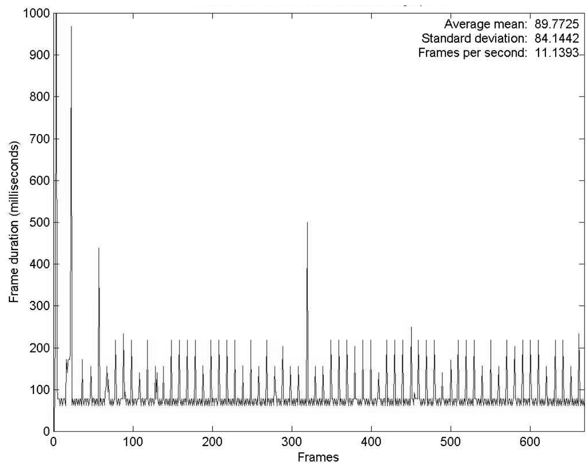


Figure 27: The World-up graph scene. Courtesy of EAI Co. Reprinted by permission

A 3-D programming toolkit which is free is Java3D produced by Sun Microsystems (Palo Alto CA, USA). Java3D programming is also based on a scene graph. However, the software is still under development, and certain drawbacks exist, when compared with WTK. One of the most important limitations of Java3D is its inability to deliver a uniform rendering speed, as uncovered by recent tests done at Rutgers University. Figure 26 [Boian, 2000] shows the same scene being rendered on a dual-processor 450 MHz Pentium PC, using (a) WTK (release 8) and (b) Java 3D (release 1.1.2). The scene consisted of 40,000 textured polygons, and collision detection was activated. When WTK was used, the average time to render one frame was 123 ms (8.1 frames/sec), with a standard deviation of about 10 ms. Interestingly enough, Java3D was 37% faster, with an average rendering speed of 11.1 frames/sec. Its average time to render a frame was only 90 ms. Unfortunately, its standard deviation was 84 ms, or 840% larger that for WTK.



a

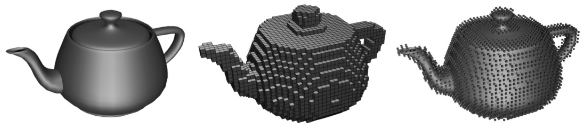


b

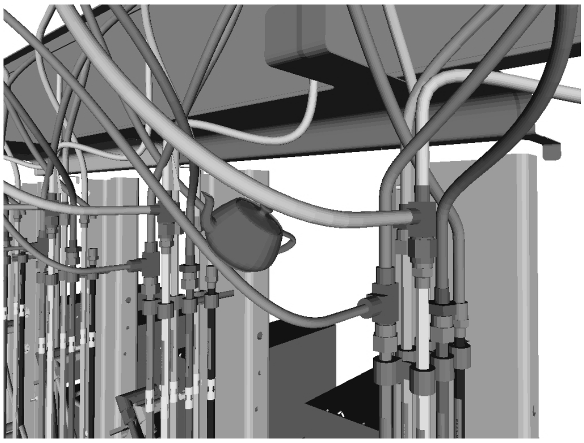
Figure 28: Comparison of frame rendering speed and consistency between: a) WTK; b) Java3D [Boian, 2000]. Reprinted by permission

Generalizations can be risky, and certainly SUN Microsystems will address some of these drawbacks in newer Java3D releases. However such large standard deviations in frame rendering time, as present in the current Java3D release will adversely impact interactions in the virtual environment, especially where force feedback is concerned.

Force feedback calculation is preceded by a collision detection step that is used by the computer to determine if there is interaction in the virtual environment. Such an algorithm needs to be both accurate and fast, which is difficult in complex virtual environments. One example is CAD analysis for accessibility. Complex assemblies, such as "crowded" aircraft engines, are difficult to design and even more difficult to service. Researchers at Boeing Co. (Seattle WA, USA) have developed the "voxel point shell" (VPS) method of collision detection to cope for such application needs [McNeely et al., 1999]. VPS builds a point shell around the surface of a single moving object in a pre-computing stage. At run time, this point shell is checked for collision with the static environment, and the resulting force/torque applied to the user. Tests done using a complex model of a Boeing 777 with almost 600 thousand polygons, shown in Figure 29, allowed haptic rendering at a constant rate of 1000 Hz. The visual frame rate was 20 frames/sec, using Boeing's proprietary "FlyThru" rendering software.



a



b

Figure 29:

6.2 Special-purpose modeling toolkits

Special-purpose toolkits have been developed to help certain types of simulations. For example, Virtual Technologies have introduced the VirtualHand® Suite 2000, which is a library designed to work with the CyberGlove, CyberGrasp, and CyberTouch interfaces [Virtual Technologies, 2000]. It helps develop applications where interaction with the objects is at the level of the hand, and includes collision detection, a force feedback API and networking capabilities.

Another special-purpose toolkit is the GHOST library developed by SensAble Technologies for their PHANToM arm. It allows the mixing of scene graph and direct force field programming, in scenes with complexities up to 250,000 polygons (mesh configuration). Multiple PHANToM Desktop models can be supported in a daisy-chain arrangement on a single host communication port.

Finally, the DI-Guy library developed by Boston Dynamics (Cambridge MA, USA) helps program simulations involving dismounted infantry, special operations and peacekeeping operation tasks by providing an intelligent-agent based library [Boston Dynamics Inc., 1997]. As can be seen in Figure 30, the toolkit allows users to control avatars that respond to real-time task-level control. Once they are given behavior (walk, kneel, crawl, etc.) and travel parameters, they execute the action through motion interpolation. This allows multiple DI-Guy characters to be included in a given virtual scene. The toolkit is currently supported by WTK (Release 9) and by Vega (Paradigm Simulations Inc., Dallas TX, USA). Vega LynX allows a point-and-click interaction environment.

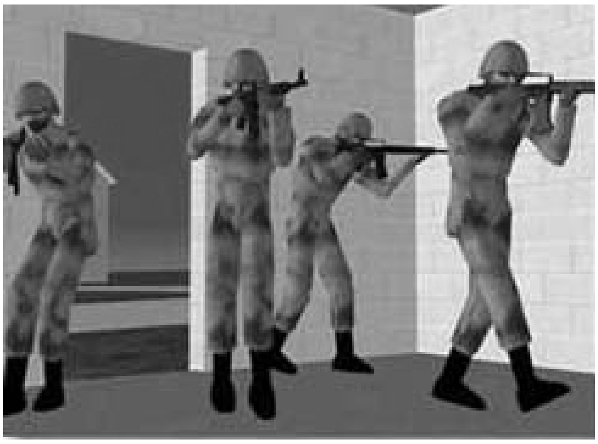


Figure 30: Scene created with the DI-Guy toolkit for dismounted infantry training. Courtesy of Boston Dynamics Inc. Reprinted by permission

7. Conclusions

There is no doubt that VR technology has been going through a rapid change. A major impact on the widespread use of this technology in the military and other areas is the tremendous decrease in computer prices, and increase in PC-based graphics speed. The miniaturization of the PC in its present form allows for portability, which results in increased user freedom of motion and simulation realism. Large-volume displays are also adding to the user ability to interact with large simulation volumes. New trackers have overcome the limitation of magnetic technology and can be used for wide area tracking and interaction. Portable haptic interfaces also add to realism, especially in tasks involving manual dexterity. Programming toolkits now offer a complex programming environment integrating the various modalities of interacting with the virtual world. All these developments point to more useful military application of VR, primarily in training, but also in C&C and weapon design/prototyping. Human factor studies need to validate the technology and its usefulness.

Acknowledgements

NATO travel support for delivery of the Key note presentation at the Workshop in The Hague is gratefully acknowledged. Author's research reported here was supported by grants from the National Science Foundation, from Office of Naval Research (DURIP) and from Rutgers University (SROA and CAIP grants).

References

Boian, R., "A Comparison Between WorldToolKit and Java3D," Rutgers University, ECE Dept., Project Report, May 2000.

Boston Dynamics Inc., "Vega DI-Guy™," Data sheet, 1997. Also at www.bdi.com.

Burdea, G. & Coiffet, P., *Virtual Reality Technology*, John Wiley & Sons, New York, 1994.

Burdea, G., *Force and Touch Feedback for Virtual Reality*, John Wiley & Sons, New York, 1996.

Dimension Technologies, Inc. "3-D Flat Panel Virtual Window Display Family," Company brochure, 4 pp., Rochester, NY, 1998.

Hix, D., Swan, E., Gabbard, J., McGee, M., Durbin, J. & King, T., "User-Centered Design and Evaluation of a Real-Time Battlefield Visualization Virtual Environment," *Proceedings of IEEE Virtual Reality'99*, pp. 96-103, March 1999.

Hollerbach, J., Thompson, W. & Shirley, P., "The Convergence of Robotics, Vision, and Computer Graphics for User Interaction," *The International Journal of Robotics Research*, vol. 18, no. 11, pp. 1088–1100, November 1999.

InterSense Co., "InterSense IS-900 Precision Motion Tracker," Company brochure, Burlington, MA, 2000. Also at www.isense.com.

Isdale, J., "Alternative I/O Technologies," *VR News*, Vol. 9, No. 2, pp. 24-29, March 2000.

Jense, H., Personal communication, TNO Physics and Electronics Laboratory, The Hague, The Netherlands, August 1993.

Kramer, J., Lindener, P. & George, W., "Communication System for the Deaf, Deaf-Blind, or Non-Vocal Individuals Using Instrumented Glove," US Patent 5,047,952, September 10, 1991.

Kramer, J., "The Haptic Interfaces of the Next Decade," Panel Session, *IEEE Virtual Reality 2000 Conference*, March 2000.

Marcus, M., "Practical Aspects of Motion Capture Technology for the Entertainment Industries," Mirage Virtual News Bulletin, www.itc.co.uk/mirage, 1997.

McNeely, W., Puterbaugh, K. & Troy, J. "Six Degree-of-Freedom Haptic Rendering Using Voxel Sampling," *Computer Graphics Proceedings* (SIGGRAPH), pp. 401-408, August 1999.

Medl, A., Marsic, I., Andre, M., Liang, Y., Shaikh, A., Burdea, G., Wilder, J., Kulikowski, C. & Flanagan, J., "Multimodal Man-Machine Interface for Mission Planning", *Intelligent Environments — AAAI Spring Symposium*, March 23-25, Stanford University, Stanford CA, pp. 41-47, 1998.

Noma, H., Sughihara, T. & Miyasato, T., "Development of Ground Surface Simulator for Tel-E-Merge System", *Proceedings of IEEE Virtual Reality 2000,* IEEE, 2000, pp. 217-224.

Olympus Corporation of America, "Eye-Trek Specifications," http://www.eyetrek.com, 2000.

Systems Technology Inc., "Parachute Flight Training Simulator," http://www.systemstech.com/paramain.htm, 2000.

Real Time Graphics, "High-Performance Image Generators — A Survey," vol. 8, no. 6, January 2000.

Trefftz, H. & Burdea, G., "Calibration Errors in Large-Volume Virtual Environments," CAIP TR-243, Rutgers University, 2000.

Virtual Technologies Inc. (2000). "VirtualHand® Suite 2000", Palo Alto CA, www.virtex.com/products.

Simulating Haptic Information with Haptic Illusions in Virtual Environments

Anatole Lécuyer[1]
Aerospatiale Matra CCR
12 rue Pasteur, 92152
Suresnes, France

Sabine Coquillart
INRIA Rocquencourt
Domaine de Voluceau, 78153
Le Chesnay Cedex, France

Philippe Coiffet
LRP – CNRS
10-12 avenue de l'Europe, 78140
Vélizy Villacoublay, France

Abstract

This paper presents a set of experiments in which a human user feels haptic sensations. These sensations are in fact haptic illusions, generated by a visual effect. Then, these haptic illusions are described and analysed.
These haptic illusions were generated by the use of a pseudo-haptic feedback system. It is a system combining an isometric input device and visual feedback. The experimental apparatus did not use any force feedback interface.
The paper addresses the role of action in the perception loop — subjects felt a reactive force corresponding to their own sensory-motor command. In addition, subjects had to "participate" in the illusion process by choosing the cognitive strategy, which led to the illusion.
In the future, the use of the concept of illusion might improve or simplify VR simulations and pave the way to a better understanding of human perception.

1. Introduction

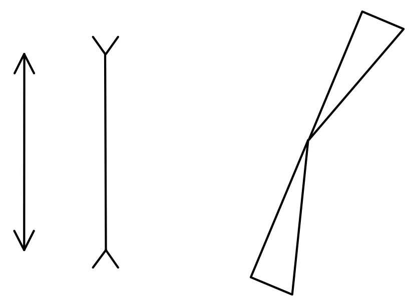
The challenge of VR technology applied to aeronautical virtual prototyping is the backdrop to the study. Nowadays, virtual prototype designers should take into consideration the assembly and the support constraints as early as possible in the development process. Indeed, the operator (or the designer) should have the possibility to feel and interact more physically with the mock up. It is therefore essential to allow haptic feedback in virtual assembly and support operation simulations.
Haptic feedback devices will soon provide new and indispensable possibilities [4]. But today these interfaces remain expensive and complex. Thus, there is a need for other replacement solutions.
In the absence of a haptic interface, a previous paper [10] studied the possibility to simulate force cues with an input device within a virtual environment (VE). This device is the 2003C model of the Logitech Spaceball™ [3], which is an isometric device — "isometric" meaning that the Spaceball™ is nearly static and remains in place while a pressure is being exerted upon it. The force feedback was simulated by using the mechanical characteristics of the passive device: its internal stiffness and its thrust — and by combining them with an appropriate visual feedback. The result of this visio-haptic feedback was called "pseudo-haptic" feedback [10]. The pseudo-haptic feedback was established qualitatively and quantitatively following different psychophysical experiments.

The simulation of force feedback by pseudo-haptic feedback can be considered as a phenomenon of haptic illusion. An illusion is a non-veridical perception. It is a mistake made by our brain and not by our senses. The effect of illusion can be generated by means of art, artefact or special effects. This effect can be perceived but is not real.
The objective of the study is to analyse some haptic illusions involved in the pseudo-haptic feedback, in order to introduce the concept of illusion in the design of virtual environments.

After addressing previous work on haptic illusion, this paper describes two different experiments, which were carried out to demonstrate the potential of pseudo-haptic feedback. Then the paper studies the haptic illusions, which are generated by these experiments. Finally, it assesses the perceptual mechanisms involved in the process as well as their potential.

2. Previous Work

Some well-known optical illusions such as the Müller-Lyer illusion (see Figure 1a) or the Zollner illusion are extensively described in scientific works [8]. Many examples of famous illusions can be found on the web [1]. And there are even companies whose business is devoted to developing educational and fun products relating to visual illusions [2].



a **b**

Figure 1: Müller-Lyer Illusion : the left segment looks smaller than the right one [**a**]; Bourdon Illusion : the left border looks slightly bent [**b**]

But illusions may occur on the other sensorial modes. An auditory illusion [2], composed by Roger Shepard in

[1] For contact with authors: anatole.lecuyer@aeromatra.com, sabine.coquillart@inria.fr, and philippe.coiffet@robot.uvsq.fr

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

1964, is a transposition of the famous endless stairs drawn by the Dutch graphist M.C. Escher on the auditory mode. Shepard played on a keyboard an ascending or descending chromatic or diatonic scale using four parallel octaves simultaneously. The tones were perceived as continuously increasing or decreasing in pitch, however after travelling over an octave, they were the same in pitch as when first started.

The existence of haptic illusions can be revealed by simple experiments. For example, considering three jars of water; from left to right, the temperature varies from warm, tepid to cold. When the hands are first dipped into the outer jars the water is perceived as warm on the left hand-side and as cold on the right hand-side. Then, when dipping both hands into the middle container, one perceives again two different temperatures, this time however, in reversed order cold on the left and warm on the right, though the water is neither warm nor cold but tepid.

The Thaler haptic illusion can also be simulated very simply. One can observe that the temperature of an object influences the haptic perception of its weight: a cold coin seems heavier than a coin of the same size but warmer [12]. Thanks, probably, to the fact that perception of coldness and that of heaviness share common neurones.

Another example described in [5], is the haptic equivalent of the Bourdon's visual illusion (see Figure 1b). Day used a 3D volumetric model of the Bourdon Figure. When a person explores the two opposite surfaces of the model with his/her thumb and forefinger, he/she feels the upper and straight surface as being slightly bent. On average, people felt a bend of 3.8 degrees *visually* and a bend of 3.5 degrees *haptically*.

Researchers do not always agree on what causes illusions, and many illusions remain "unsolved". Ellis and Lederman focused more precisely on the origin of illusion as located on the visual mode or the haptic one. They studied the famous size-weight illusion [7] and the material-weight illusion [6] — the size-weight illusion occurs when a large radius ball seems heavier than a ball of the same weight but with a smaller radius. The material-weight illusion is the influence of the texture of an object on the perception of its weight. Ellis and Lederman established these two illusions as a primarily haptic phenomenon, despite the size-weight illusion was traditionally considered as a case of vision influencing haptic processing.

Some works deal with consequences of illusions on perception or on the performances of our motor system. Volker studied the influence of visual illusions on grasping [15]. Different subjects were presented fins on a monitor screen being directed either outwards or inwards such as in the Müller-Lyer illusion (Figure 1a). During the grasping task, subjects were told to grasp the fin, and the maximal aperture between thumb and index finger was measured. During the perception task, subjects were told to adjust the length of a comparison bar on the screen to match the length of the fin. Volker showed that there were strong effects of the Müller-Lyer illusion on grasping as well as on visual perception, indicating that the motor system is also receptive to visual illusions.

In a VR simulation, Hogan studied haptic illusions occurring during the exploration of virtual objects and their implications on the perceptual representation of these objects [9]. He used a force feedback arm constrained to move on the 2D horizontal plane. Subjects grasped the handle of the arm and were asked to evaluate the length and stiffness of virtual rectangle objects. The handle was grasped and moved around the virtual rectangles. During the task, the subject had to choose the longer (or the stiffer) stimulus of two stimuli. One stimulus was evaluated on the X-axis of the horizontal plane, while the other was evaluated on the Y-axis (i.e. each stimulus depended on a side of the rectangle). Results show that for the same stimulus on the X- and Y-axis, a difference of perception occurs according to the distance of the handle from the shoulder (as if the vertical-horizontal visual illusion was projected in the haptic mode, on the horizontal plane). Hogan stated that these haptic illusions show that the internal model of haptic perception is not metrically consistent. This property should significantly modify and simplify the performance constraints in forces computation.

It seems that very few VR papers explored the possibility to use illusions directly in the conception of a VE.

This paper presents and analyses haptic illusions, which were showed by two VR experiments. The next part describes the two experiments and their results.

3. Pseudo-Haptic Experiments

The concept of pseudo-haptic feedback relies on coupling the visual feedback with the internal resistance of the isometric device, which naturally reacts to the force applied by the user. The overall system returns a force information called pseudo-haptic feedback.

For example, let us assume that an operator manipulates a virtual pipe in a virtual environment within the frame of an insertion task evaluation. The pipe is displayed on the monitor, and moved by means of the Spaceball™. It is to be inserted into a virtual duct. As the pipe penetrates the duct, its speed is slowed down. The user instinctively increases his pressure on the ball, which results in the feeding back of an increasing reaction force by the static device. This combination of visual effect and growing reactive force is then expected to generate cues of friction.

In order to study the pseudo-haptic feedback concept, different experiments were conducted. Two of them are described in the following paragraphs.

3.1 The "Swamp" Experiment

Description

The swamp is a quantitative evaluation of the pseudo-haptic feedback. 18 people took part in this experiment. Each subject was told to manipulate a virtual cube in a

3D virtual environment (see Figure 2). The cube was manipulated in 2D on the horizontal plane with either a classical 2D mouse or with the Spaceball™.
As the cube moves over a grey area, its speed is accelerated or slowed down. At this very moment, the subjects were asked to describe and compare their sensations when using the 2D mouse or the Spaceball™.

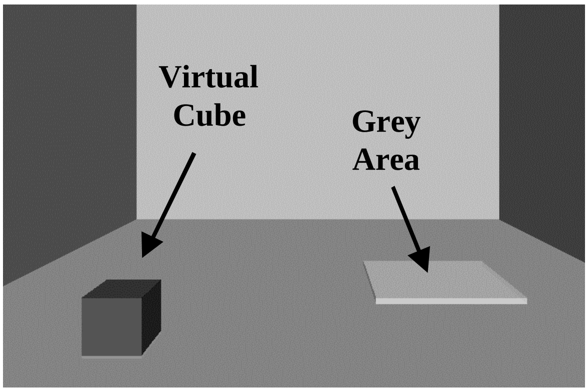


Figure 2: The Swamp experiment display

Results
A quantitative comparison between an isometric device and an isotonic device must be taken cautiously since these interfaces are not used in the same way. But the swamp example *did* display some global tendencies.
A great majority of people logically found that the use of the two interfaces were very different. The need for a learning phase with the Spaceball™ generally disturbed subjects when starting their manipulation.
The subjects systematically perceived the following phenomena: *friction*, *gravity* or *viscosity,* when the cube was slowed down with both devices. Conversely, they perceived a sense of *gliding* or *lightness* when the cube was accelerated.

Great majorities of people found that the sensations they felt were different while using the Spaceball™ or the 2D mouse. Nearly all of the subjects chose the Spaceball™ as the interface with which the "forces" were more perceptible. This sensation was less obvious when the cube was accelerated — which is probably due to the fact that the reactive force from the static device is more efficient during a compression phase.
The quantitative indications provided by the swamp experiment were very useful to show us the potential of this concept, but they didn't measure the characteristics of the generated feedback. It was necessary to evaluate more qualitatively the pseudo-haptic information: to do so a psychophysical experiment has been conducted

3.2 Discrimination between a Virtual Spring and a Real One
Description
The psychophysical task, which was chosen, is manual compliance discrimination between a virtual spring and a real one (see Figure 3). The real spring is embedded inside a piston, like a "trumpet piston" (see Figure 3).

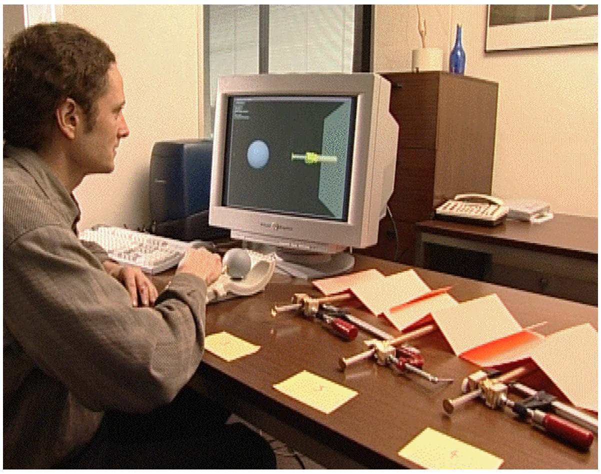


Figure 3: Psychophysical experiment — Manual discrimination between a virtual spring and a real one

The virtual spring is a combination of the input device and the visual feedback (see Figure 4). A hand-made apparatus was fixed on the Spaceball™ to obtain the same catching in the virtual environment and in the real one. The virtual spring is visually displayed on the computer screen. It is made to appear as similar as possible to the real piston. The force applied on the ball by the user controls the visual displacement of the virtual spring. When pressing the virtual spring, the user's thumb barely moves, since the Spaceball™ is an isometric — hence static — device.

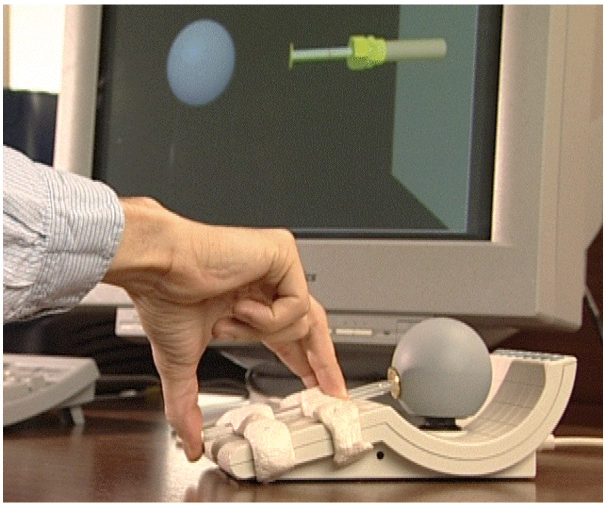


Figure 4: Virtual spring set-up

27 people took part in this experiment. There were 972 trials per subject. During each trial, each subject was asked to test a real spring and a virtual one and to select the one, which seems to him to be the stiffer. There were three possible real springs with three different degrees of stiffness. And each real spring was compared with 12 different virtual springs.

Theoretically, a “stiffer” virtual spring corresponds to a case when the force — which is required to move the visual display of the piston on the screen along a certain distance — is bigger than the one which is required to move the real spring along the same distance.
(For a complete description of this experiment see [10].)

Results
The large volume of collected data made it possible to calculate a psychophysical parameter called the Just Noticeable Difference (JND). The resulting average JND for the manual compliance discrimination between a virtual spring and a real one is equal to 13.4%. It is consistent with previous studies on compliance discrimination between two springs simulated within a single environment [13].

This consistency shows quantitatively that a system, which combines visual feedback and an isometric device, can provide force cues, which are comparable with real ones.

4. Illusions Observed

The whole concept of pseudo-haptic feedback relies on a phenomenon of haptic illusion. In the course of the two experiments, the haptic perception is mistaken by a visual effect. The visual feedback generates a new haptic interpretation of a virtual scene, thus a haptic illusion.

This assumption is confirmed by the simple fact that if one closes one’s eyes during one of these two experiments, the experimental task becomes impossible, and the haptic sensations vanish.

During the first experiment, the perception of friction when crossing the grey area in the virtual environment is linked to the visual variation of the speed of the cube. The whole set-up generates a haptic illusion of several haptic attributes of the cube — heaviness, lightness — or of the grey area — rugosity, viscosity, and friction.

In the course of the second experiment, without the visual displacement the haptic perception of the virtual spring remains the same, i.e. the Spaceball™ internal stiffness. The pseudo-haptic set-up generates the haptic illusion that different springs are being manipulated. It becomes possible to perceive different stiffness with the same Spaceball™.

One more illusion phenomenon is revealed in the course of the second experiment by a question asked to the last ten subjects. These people were told to draw a straight line corresponding to the maximum displacement of the thumb when pressing a virtual spring. The result indicates an average overestimation of 5 times their actual displacement (see Figure 5). It means that they completely assimilated the visual displacement on the computer screen to their own thumb motion. In other terms, it implies an illusion of their proprioceptive sense.

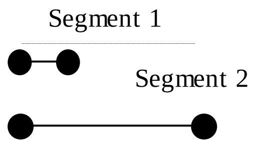


Figure 5: Illusion of the Proprioceptive Sense. Segment 1 — real maximum displacement of the user’s thumb; Segment 2 — estimated displacement of the user’s thumb

5. Discussion

The pseudo-haptic feedback is not an illusion of force feedback. There is actually a force feedback during both experiments when actuating the Spaceball™:

First, because in all manipulation tasks there is a force feedback reaching the brain — in terms of pain or fatigue for example. Broadly speaking, even when one simply holds something in one’s hand or wants to grasp an object, the motion command sent from one’s brain activates the muscles of the arm, and makes the one feel efforts or tensions in his/her muscles or his/her tendons. These efforts being sent back to the brain via the afferent neurone network.
The manipulation of the virtual cube with the 2D mouse in the first experiment could then be considered as a case of pseudo-haptic feedback with an isotonic device. The speed of the cube was decreased when passing over the grey area, then the user had to increase his arm motion, spend more energy on this gesture, and this may lead to the friction effect.
In addition to the afferent signals coming from the different mechanoreceptors, some efferent mechanisms play a role in human kinaesthesia. Such as the “innervating sensation” [14], which occurs when one overestimates the weight of an object when tired. It is a distortion of force perception, which is due to our own command system. Our will to achieve an action generally makes us feel the anticipated result of this action before it actually happens. And this introduces the role of action in the perception loop of illusion.

Then, a force feedback from the static device is also present. It is not an “active” force feedback — i.e. a computed force feedback — different from other force feedback systems such as the PHANToM™ [11] of SensAble Technologies. The current force feedback is provided by the reactive force coming from the Spaceball™. And since the Spaceball™ is nearly static, it means that the reactive force is nearly equal to the force applied by the user on the ball. It is a characteristic of the pseudo-haptic feedback with an isometric device: the force feedback is always equal to the force command. This is illustrated on Figure 7.

Figure 6 and Figure 7 show the difference of information flux in a pseudo-haptic system and a haptic one. In the case of a classical haptic feedback system such as the PHANToM™ (see Figure 6), the user transmits a motion

which is sensed by the optical encoders of the PHANToM™. The force feedback device sends back to him the computed virtual force, which is a function of the interference between the probe and the virtual elements of the virtual environment. The visual feedback doesn't play a major role in the haptic perception process.

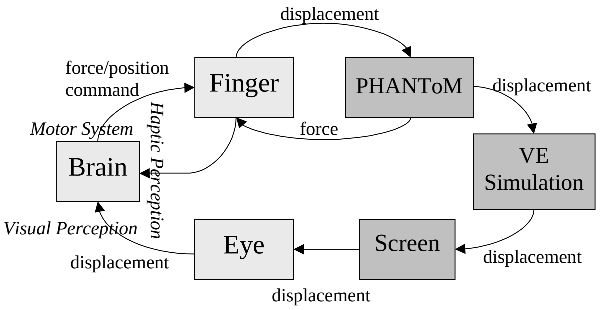


Figure 6: Haptic feedback system

In the pseudo-haptic case (see Figure 6), the user transmits his force command to the simulation by means of the force sensors of the Spaceball™. At each simulation step, the force fed back is constantly equal to the opposite of the force applied. The user receives exactly the same force as the one he has just applied. The haptic profile of the force vs. displacement is given by the Spaceball™ internal stiffness. This profile corresponds to the one of a constraint gauge; this profile is not linear. The final haptic perception is achieved by combining the force information and the visual impact of this information in the VE. It means that the Spaceball™ internal stiffness is somehow "mapped" on a visual event. In reverse, the visual effect gives sense to the force information.

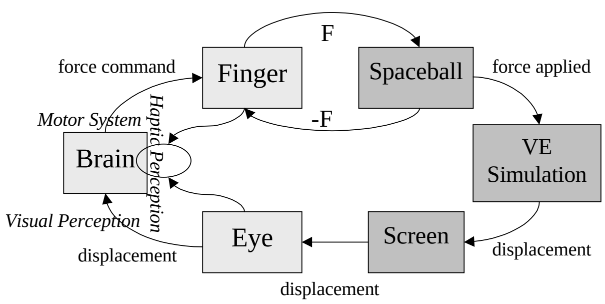


Figure 7: Pseudo-haptic feedback system

An obvious consequence of this characteristic being that the pseudo-haptic experiments, which were described, cannot work without visual feedback.

The pseudo-haptic feedback process system implies that the user receives his/her own force command in return. Indeed, the whole pseudo-haptic feedback depends on an action as well as a participation of the user during simulation: in the course of the swamp experiment, the friction sensation occurs if the user increases his pressure on the ball when the cube motion is slowed down. This probably happens if he/she decides to keep the virtual cube at fast motion, which is a cognitive strategy relying on many factors affecting the subject (passivity, availability, stress, etc.). This would imply that pseudo-haptic feedback and haptic illusion could also depend on cultural or contextual reactions of the subject.

In the course of the compliance discrimination task, the subject had to recompose the stiffness of a virtual spring with information coming from different modalities. In addition, there was a conflict concerning the spring displacement between the proprioceptive information and the visual one.

Since they were able to compare the final model of the virtual spring with a real one, the result of the experiment shows that subjects succeeded in recombining sensory information. It implies that they made the choice to use the visual displacement rather than the proprioceptive displacement. This choice is the result of an unconscious participation of the user in pseudo-haptic simulation, and is the reason why the illusion appeared.

For the time being it is difficult to know if this choice is:

- an example of a sensory substitution or sensory dominance, which corresponds to the following expression: *Vision dominates Touch $\Rightarrow$ I use the visual displacement to evaluate virtual springs,*
- or, an example of a choice between different cognitive strategies. I must choose among all possibilities one that can help me to perform my discrimination task, and eliminate other strategies. This rather corresponds to the second expression: *I must evaluate a virtual spring $\Rightarrow$ I choose the visual displacement (and not the proprioceptive one) which makes it possible.*

In other terms, is the proprioceptive illusion due to a characteristic (or a limit) of human perception system (= "peripheral" view); or is it due to a decision process made in a strategic situation (= "central" view).

This alternative has a direct impact on the conception of VE's which are to be based on pseudo-haptic feedback or sensory illusions. There is indeed, a need for further investigation concerning the generation of illusions.

6. Conclusion

The paper has presented VR simulations in which force cues or haptic behaviours are simulated with a pseudo-haptic feedback. This pseudo-haptic feedback comes along with phenomena of haptic illusions. It is not an illusion of force feedback, but rather an illusion of using a force-feedback device.

The analysis of a pseudo-haptic feedback system shows the role of the sensory motor command in the perception loop, and also points to the unconscious participation of the user in the illusion, which is linked, to his/her cognitive strategy during the experimental task.

Designers of virtual environments, who usually try to recreate human stimuli in an anthropomorphic manner, could envisage a wider use of this concept of illusion.

The method, should it exist, implies to revise the simulation process and the use of human-computer interfaces. The designer has to think in terms of sensory information feedback. He/she has to decompose the sensory information into its different sensory modalities, and to reshape it into a new sensory distribution. To do so, he/she can make full use of all the possibilities that are known in the field of sensory illusions and sensory substitutions.

It is necessary to facilitate the repositioning of the user perception to an "implicit" solution. It means that this implicit sensory alternative must be explicit enough to be found quickly by the user.

For example, in the case of the second experiment, the information needed was the displacement of the virtual spring, and its implicit alternative was the visual displacement.

Future work must develop and evaluate more cases in which sensory illusions are used for VE interactions. The overall objective is to propose an empirical method to incorporate illusions in the conception of VE's.

Acknowledgements

The authors would like to thank Mr P.R. Persiaux, Mr. D. Tonnesen and Mrs. M.J. Paškauskaitė for their valuable remarks.

References

[1] http://www-psy.ucsd.edu/~sanstis/SASlides.html

[2] http://www.illusionworks.com

[3] http://www.spacetec.com

[4] G. Burdea. *Force and Touch Feedback for Virtual Reality*. John Wiley and Son, US, 1996

[5] R.H. Day. The Bourdon Illusion in haptic space. *Perception and Psychophysics*, 47, 400-404, 1990

[6] R.R. Ellis and S. J. Lederman. Modality, Weight and Grip Force Effects in the Material-Weight Illusion. In *Proc. of the Canadian Society for Brain; Behavior and Cognitive Science Annual Meeting* , 1995

[7] R.R. Ellis and S. J. Lederman. The Role of Haptic versus Visual Volume Cues in the Size-Weight Illusion. *Perception and Psychophysics*, 53(3): 315-324, 1993

[8] E.B. Goldstein. *Sensation and Perception*. Brooks/Cole, US, 1999

[9] N. Hogan, B.A. Kay, E.D. Fasse, and F.A. Mussa-Ivaldi. Haptic Illusions : Experiments on Human Manipulation and Perception of "Virtual Objects". *Cold Spring Harbor Symposia on Quantitative Biology*, 55:925:931, 1990

[10] A. Lécuyer, S. Coquillart, A. Kheddar, P. Richard, and P. Coiffet. Pseudo-Haptic Feedback : Can Isometric Input Devices Simulate Force Feedback? In *Proc. of IEEE International Conference on Virtual Reality*, 2000

[11] T.H. Massie, J.K. Salisbury. The PHANTOM Haptic Interface : A Device for Probing Virtual Objects. In *Proc. of ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, 1994

[12] Sherrick and Cholewiak. A Finite Element Formulation for Nonlinear Incompressible Elastic and Inelastic Analysis. *Computers and Structures*, 26(1/2):357-409.

[13] H.Z. Tan, N.I. Durlach, G.L. Beauregard, and M.A. Srinivasan. Manual Discrimination of Compliance Using Active Pinch Grasp: the Roles of Force and Work Cues. *Perception and Psychophysics*. 57(4):495-510, 1995

[14] C. Tzafestas. *Synthèse de retour kinesthésique et perception haptique lors de tâches de manipulation*. Ph.D. Thesis, Université de Paris 6, Jul. 1998

[15] F. Volker, M. Fahle, K.R. Gegenfurtner, and H.H. Bülthoff. Grasping visual illusions: No difference between perception and action? *In Proc. of ARVO Meeting*, 1999

Tactile Displays in Virtual Environments

Jan B.F. van Erp[1]
TNO Human Factors
Kampweg 5
3769 DE Soesterberg
The Netherlands

Summary

Virtual Reality (VR) technology allows the user to perceive and experience sensory contact with a non-physical world. A complete Virtual Environment (VE) will provide this contact in all sensory modalities. However, even state-off-the-art VEs are often restricted to the visual modality only. The use of the tactile modality might not only result in an increased immersion, but may also enhance performance. An example that will be discussed in this paper is the use of the tactile channel to support the processing of degraded visual information. The lack of a wide visual field of view in VEs excludes the use of peripheral vision and may therefore degrade navigation, orientation, motion perception, and object detection. However, tactile actuators applied to the torso have a 360° horizontal 'field of touch', and may be suited to present navigation information.

1. Introduction

Developments in VR technology have mainly focussed on the visual sense. In the last decade, enormous improvements have been made regarding the speed and resolution of the image generators. However, the human senses are not restricted to the visual modality. Using the auditive and tactile modality as well in a VE might have several advantages. This paper will more specifically discuss the tactile sense in relation to VE use. I will restrict the tactile channel to 'the skin as information channel'. Thus, I will not include receptors in muscles and joints as part of the tactile sense. When these are included, one usually uses the term haptics. On the other hand, tactile information is not restricted to 'touching' (i.e., feeling objects), but also comprises (passive) vibro-tactile stimulation of the skin and temperature perception.

Employing the tactile modality has several potentially useful applications and advantages in VE, including the following:

1. The quality of the VE and user performance is likely to improve if the information that is available to the tactile sense in real life is present in the VE as well. This is certainly true for information that is predominantly perceived with the tactile channel, such as roughness of objects, and small vibrations.
2. Employing the tactile sense will enlarge the immersion of the observer in the VE. The VE is more complete, and sensory information may become congruent: I can feel what I see.
3. Tactile information can guide movements. An example is the potential role of tactile information in grasping. Users may have trouble in estimating the distance between their (virtual) hand and the object they want to grasp. Presenting a tactile gradient (i.e. a tactile intensity or frequency field around the object) which guides the user to the object and indicates the Euclidian distance between the object and the user's hand might support the degraded visual information in VEs. After grasping the object, tactile information may be used to indicate how much force must be applied to the object (see next point).
4. Tactile information can be a substitute for force feedback. Force feedback is essential for adequate user performance in interacting with virtual objects (e.g., instruments and weapons), but is also very difficult to present with contemporary VR technology. Tactile information as a substitution for force feedback has already proven its effectiveness in remote control situations.
5. The tactile sense may be helpful in overcoming the weak points that even state-of-the-art VE systems still have. For example, the field of view of the visuals is still reduced compared to real life; using the tactile sense to compensate for the lack of peripheral viewing is one of the possibilities.
6. Finally, the tactile modality may be used as a general information channel to present VE-related but not specific information, e.g., warning information.

For all these applications fundamental and applied knowledge is required for successful use in VEs, and moreover, for successful development of devices. At this moment, not all this knowledge is available or applicable. Areas that deserve attention include:

- body loci other than hand and fingers,
- sensory congruency (below, an example shows that this doesn't come naturally),
- cross-modal interaction,
- perceptual illusions,
- attention.

A simple experiment by Werkhoven and Van Erp (1998) showed that visual and tactile information is not always perceived consistently. They investigated the perception of open time intervals, either marked by visual stimuli (blinking squares on a monitor) or tactile stimuli (bursts

[1] For correspondence with the author: vanerp@tm.tno.nl

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

of vibration on the fingertip. They compared standard intervals of 200 ms with uni- and cross-modal intervals, as is schematically presented in Figure 1 for the cross-modal condition.
The results of this experiment showed a large bias in the cross-modal condition: tactile time intervals are overestimated by 30% (see Figure 2). This indicates that sensory congruency is a non-trivial aspect of integrating sensory modalities in a VE.

Overview of the paper
This paper focuses on the use of the tactile modality to present navigation (i.e., direction) information. This application can help VE users in orientating in VEs, which may be difficult on the basis of restricted visual information only.
In the next section of the introduction, some examples of tactile displays are given. Chapter 2 describes some basic neurophysiology and psychophysical knowledge. An example of cataloguing spatio-temporal characteristics is given in chapter 3. Here, the spatial characteristics of the torso are described, including experimental data. This cataloguing is of primary interest for the application that is described in Chapter 4: using the torso to present tactile navigation information. The torso has three important advantages in this respect. First, it has a large surface, reducing the need to minimise actuator size or to keep the number of actuators low. Second, information presented to the torso does not interfere with actions performed with the hands, like controlling input devices. And third, the torso is a volume, and thus *a priori* interesting for presenting 2D or 3D information, like geographical or navigational information.

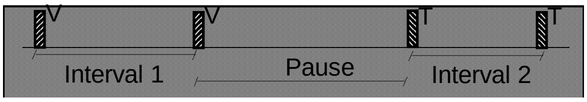


Figure 1: Schematic presentation of the stimuli to investigate the perception of open time intervals. The intervals are marked by visual stimuli (marked V) or tactile stimuli (marked T)

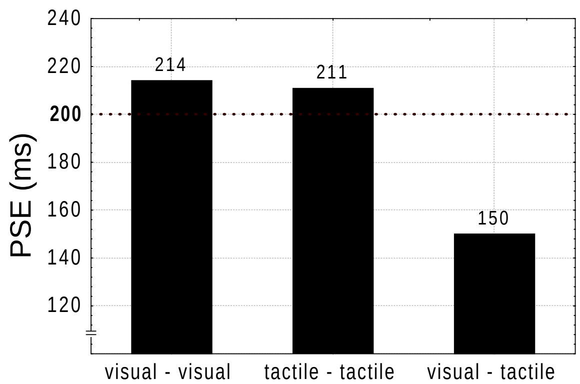


Figure 2: Point of subjective equality for a 200 ms standard open time interval experiment. The visual — tactile condition shows that a 150 ms tactile interval is judged to be equal in length to a 200 ms visual interval

Examples of tactile displays
This section gives a small and far from complete overview of tactile display applications (see also Van Erp & Van den Dobbelsteen, 1998). It focuses on two application areas: that of sensory substitution and navigation displays. This restriction is made, because displays developed for use in VE are regularly described in the open literature, e.g. see Boman (1995) or Ziegler (1996).

Sensory substitution
Some examples of the earliest displays providing complex stimuli are aids for the blind, including miniature matrices of point stimuli used for reading of text and pictures. ‘Tactile imaging’ is the process of turning a visual item, such as a picture, into a touchable version of the image, so that this tactile rendition faithfully represents the original information.

- *The Optacon.* One of the most successful devices to present ‘visual’ information to the blind was an ink-print reading machine, the Linvill-Bliss Optacon (OPtical-to-TActile CONverter). Bliss and his associates (Linvill & Bliss, 1966; Bliss et al., 1970) developed this reading device, which converts printed materials into vibratory patterns. With the aid of a small camera containing a matrix of 6 by 24 photocells, the device converts the image electronically to a tactile display, placed on the skin of a fingertip.
- *The Kinotact.* Craig (1974) studied letter-shape perception with the aid of a 10 by 10 matrix of vibrators placed against the observer’s back. The encoding system, called ‘Kinotact’, was a 10 by 10 matrix of photocells, wired one-to-one with the vibrators. With the presentation of the tactile image of block letters, subjects learned to identify this ‘pictorial mode’ letter patterns to an average criterion of 80-90% correct in 300 trials. For related research, see also Loomis (1974), and Craig (1980).
- TVSS. Bach-Y-Rita (1972) and associates developed the Tactile Vision Substitution System (TVS system), in which a visual image picked up by a TV camera is transformed into a tactile one by means of a 20 by 20 matrix of vibrators mounted on the back of a dental chair. It was found that subjects could immediately recognise vertical, horizontal and diagonal lines. Experienced users could identify common objects and people’s faces. This is an example of a perceptual phenomenon called distal attribution, in which an event is perceived as occurring at a location other than the physical stimulation site. With self-induced camera movement, subjects use the camera as part of a perceptual organ and learn to locate the percepts subjectively in space, rather than on the skin.

Another TVS system, called the Electrophthalm, developed by Starkiewicz, Kuprianowicz and Petruczenko (1971) is more applicable to space orienta-

tion and presents a 12 by 8 tactile image to the forehead. However, TVS systems are not useful for acquiring information from 'cluttered' visual environments and are not presently useful for navigation purposes.

- Desktop tactile displays. The formerly described systems are not designed to provide computer access to the visually impaired, and are rarely used due to uncomfortable or impractical displays and inefficient information transfer (Kaczmarek & Bach-Y-Rita, 1995). An example of a new generation display, which I will call desktop tactile displays, is the Moose. This display is especially designed to provide computer access. A prototype developed by O'Modhrain and Gillispie (1997) presents a haptic representation of a screen by reflecting forces when navigating across the screen. Desktop tactile displays are nowadays widely available in the consumer electronics shops for as little as 100 US$.

Tactile navigation displays

A second important application of tactile displays is as navigation display. Gilliland and Schlegel (1994) conducted studies to explore the use of vibrotactile stimulation of the human head to inform a pilot of possible threats or other situations in the flight environment. Rupert, Guedry and Rescke (1993) developed a matrix of vibro-tactors that covers the torso of the pilot's body (http://www.accel.namrl.navy.mil/default.html). This prototype may offer a means to continuously maintain spatial orientation by providing information about aircraft acceleration and direction of motion to the pilot. Within the pitch and roll limits of their torso display (15° and 45°, respectively), the subjects could position the simulated attitude of the aircraft by the tactile cues alone. The Tactor Evaluation System (TES, Engineering Acoustics Inc.) was developed to demonstrate the use of vibrotactile information for divers in conditions of low visibility: real time navigational information (course, distance, and cross-track error) and alarm information. Five tactors were used: left and right side, back and chest, and on a wrist for miscellaneous signals (http://www.eaiinfo.com/).

2. Cataloguing spatial sensitivity

An important parameter in the design and application of tactile displays is the spatial resolution. There are two main areas involved in spatial sensitivity research: neurophysiology and psychophysics. Important determinants of spatial sensitivity are the sizes and forms of the receptive fields of the mechanoreceptors, and the representation of the body surface in the (somatosensory) cortex. This neurophysiological data is presented in Section 2.1. The psychophysical measures of spatial sensitivity used throughout the years and experimental findings are presented in Section 2.2. For a more elaborate overview, see for example Van Erp and Vogels (1998). Basic research on the spatial sensitivity of the torso for vibro-tactile stimuli (relevant for the application under study) is presented in Chapter 3.

2.1 Neurophysiology

A comprehensive overview basic neurophysiology can be found in Kandel et al. (1991). An important contribution of this research area has been the determination of the density of receptors, and the size and form of the receptive field of a single peripheral nerve fibre. Micro-neurographic recordings from nerves innervating the glabrous skin have isolated four groups of mechanoreceptive fibres (see Table 1 for an overview).

After contacting a single afferent unit, a systematic exploration of the receptive field is undertaken. Unfortunately, this technique is only applied for the human arm and hand; no data on the trunk are available. Furthermore, the technique provides information on single peripheral nerve fibres only, not on the spatial sensitivity of the cutaneous sense as a whole. Applied to the Pacinian body, the receptive field proves to be large, with poorly defined borders and a single point of maximum sensitivity. Even for the fingers, receptive fields can be in the order of several square cm (Bolanowski et al., 1988; Valbo & Johansson, 1978).

Table 1: Characteristics of the four types of mechanoreceptive fibres in the human skin

	fast adapting	slowly adapting
superficial skin	**Meissner corpuscle (RA)** • small receptive field • NP I channel, not sensitive to temperature • 10–100 Hz • temporal summation: no • spatial summation: yes • local vibration and perception of localised movement	**Merkel cell (SAI)** • small receptive field • NP III channel, sensitive to temperature • 0.4–100 Hz • temporal summation: no • spatial summation: no • tactile form and roughness
deeper tissue	**Pacinian corpuscle (PC)** • large receptive field • P-channel, very sensitive to temperature • 40–800 Hz • temporal summation: yes • spatial summation: yes • perception of external events	**Ruffini ending (SAII)** • large receptive field • NP II channel, sensitive to temperature • 15–400 Hz • temporal summation: yes • spatial summation: ? • not in glaborous skin

Besides the receptive field sizes of single afferent nerve fibres, one has also determined the receptive field sizes of the different cortical regions involved in cutaneous processing.

2.2 Psychophysics

Within psychophysics, two classic measures are applied to determine the spatial resolving power: the two-point limen (participants have to judge whether a stimulus consists of one or two points) and the error of localisation (e.g. participants judge two successive contacts as the same or different in locus). Both methods know different variants. Unfortunately, little data are

available on vibro-tactile perception and on loci other than the hand.
Weber and Vierodt did the first psychophysical research on spatial acuity in the nineteenth century. It was Weber who introduced the two-point limen and the localisation error (Weber, 1834). Mapping of the whole body revealed large differences in spatial acuity between different parts of the body. Vierodt (1870) generalised this to the 'law of mobility', which states that the two-point limen improves with the mobility of the body part.

After the work of Weber and Vierodt, little attention was given to this field until the 1960s. Weinstein (1968) measured (pressure-) thresholds of two-point discrimination and tactile point localisation on several body loci. Both thresholds were highly correlated, however. Acuity found with two-point discrimination was three to four times lower than with point localisation. Because the methods of two-point discrimination and point localisation are measures for spatial acuity and hyper acuity, respectively, the results are in accordance with data on visual acuity (e.g. see Snippe, 1991). Furthermore, Weinstein found significant effects of body locus. Lowest thresholds were found for the fingertips: 2.5 mm and 1.5 mm for two-point discrimination and point localisation, respectively. Thresholds for the trunk were approximately 40 mm and 10 mm, respectively. Sensitivity decreased from distal to proximal regions: fingers, face, feet, trunk, upper and lower extremities. Thresholds correlated with the relative size of cortical areas subserving a body part. Another important observation was that good two-point discrimination did not necessarily mean good sensitivity to pressure. Vierck and Jones (1969; Jones and Vierck, 1973) stated that the method of the two-point limen leads to an underestimation of the skin's real spatial sensitivity. They showed that the discrimination of area stimuli and length stimuli is about ten times better. In the 1970s, Loomis and Collins (1978) found comparable results when the stimulus was a gradual shift in the locus of stimulation.

Johnson and Phillips (1981) introduced alternative methods, and measured two-point thresholds, gap detection and discrimination of grating orientation for the fingertips. They found thresholds of 0.87 mm and 0.84 mm, respectively. These results show that the ability of subjects to discriminate stimuli is much finer than is indicated by the two-point threshold of Weinstein (1968).

3. Cataloguing vibro-tactile spatial resolution on the torso

Since only indirect data are available regarding the spatial resolution of the <u>torso</u> for <u>vibro</u>-tactile stimuli, basic research was needed to formulate the optimal display configuration. On the one hand, one wants to use the full information processing capacity that is available; on the other hand, one wants to keep the number of actuators to a minimum. Therefore, a concise discussion of a series of experiments is presented (for details, see Van Erp & Werkhoven, 1999).

Four male subjects (age range 28–39 years, mean 31) participated voluntarily. In the experiment, 11 vibro-tactile actuators were attached to the torso with sticky tape (see Figure 3). The participants performed a localisation task: Two stimuli were presented to the torso and the participant was asked to judge the location of the second compared to the first (left/right). The stimuli were first presented to the dorsal side of the torso, and in a second session to the frontal side. The inter stimulus interval (ISI) was varied (0 ms, 56 ms, 196 ms, and 980 ms), as was body locus within a torso side (left, middle, and right). The latter indicates the location of the standard stimulus; each standard was combined with four comparison stimuli. The responses of the subject to each standard-comparison pair were counted in proportion 'to the right' responses. These summarised data were fitted to a cumulative normal distribution, resulting in two parameters: μ (or bias) and σ (or threshold), see Figure 4.

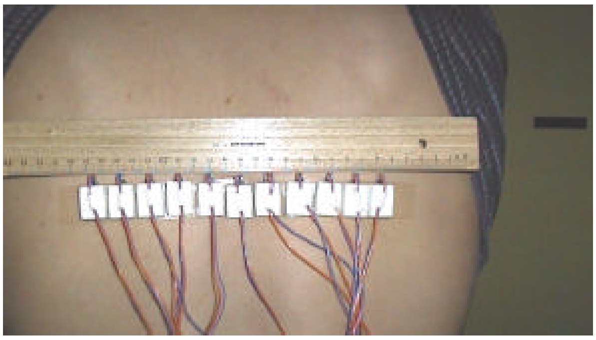


Figure 3: Placement of the tactile actuators on the back

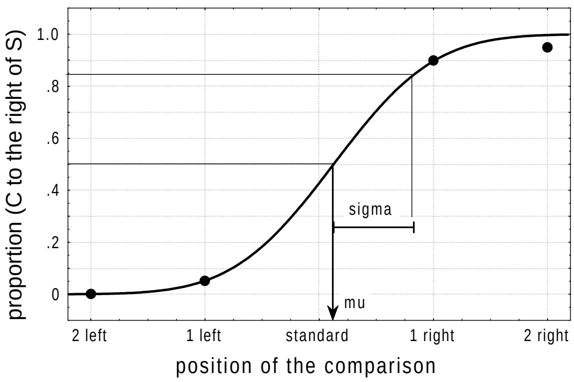


Figure 4: Psychophysical method to determine the bias (mu) and sensitivity (sigma) for a specific standard (S)

The results of the experiment (see Figure 5) showed that the sensitivity for vibro-tactile stimuli presented to the ventral part of the torso was larger than for stimuli presented to the dorsal part. Furthermore, the effect of body locus was present on both the frontal and the dorsal part: the sensitivity near the middle is larger than to the sides. Moreover, the sensitivity is larger than expected on the basis of the psychophysical literature. The effect

of ISI showed that sensitivity increases with increasing ISI.

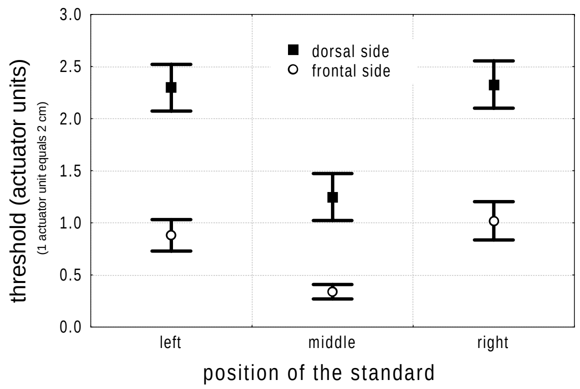


Figure 5: Results of the spatial accuracy of the torso for vibro-tactile stimuli

4. Example of implementing a tactile display: presentation of spatial information on the torso

When the first phase, cataloguing relevant perceptual characteristics, is finished, basic research into possible applications becomes actual. As discussed in the introduction, the torso may be well suited to present 2D geographical information. In the following experiment, tactile actuators were attached around the participants torso (except for the region around the spine, see also Figure 6). During the experiment, one actuator was activated. The observer could adjust a cursor to indicate the external direction suggested by the actuator (see Figure 7 for the experimental set-up).

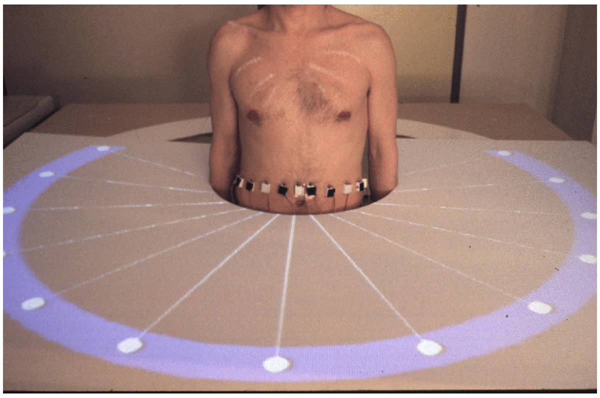


Figure 6: Method to ensure correct placement of the actuators

This direction determination task resulted in two parameters: a bias in the indicated direction, and variability in the answers (expressed in the standard deviation of the responses). The latter parameter is of course a measure of the precision with which the observer perceives the stimuli.

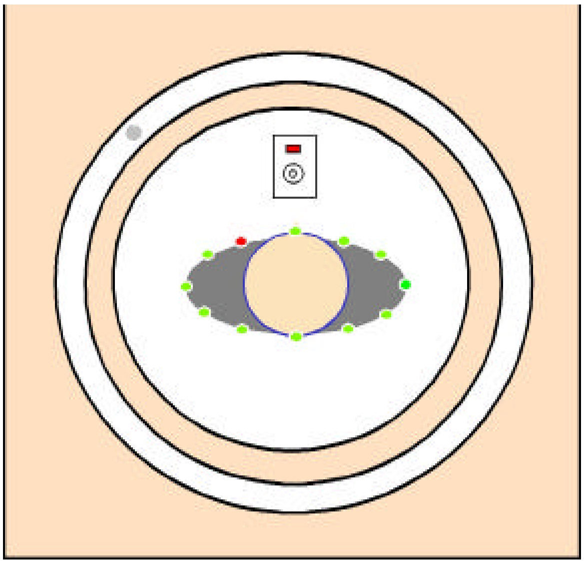


Figure 7: Top view of the set-up for the direction discrimination task. With a dial, the observer can position a cursor (a dot projected from above) along a white circle drawn on the table. The cursor should be positioned such that it indicates the direction of the tactile stimulus

The results are interesting in several ways. First of all, none of the participants had any trouble with the task. This is noteworthy since a point stimulus does not contain any explicit direction information. The strategy people use is probably equivalent to that of visual perception, namely using a perceptual ego-centre as second point. Several authors determined the visual ego-centre (e.g., Roelofs, 1959), which can be defined as the position in space at which a person experiences himself or herself to be. Identifying an ego-centre or internal reference point is important, because it co-ordinates physical space and phenomenal space. A second reason to determine the internal reference point in this tactile experiment was the striking bias all ten participants showed in their responses, namely a bias towards the sagittal plane. This means that stimuli on the frontal side of the torso were perceived as directions coming more from the navel, and stimuli on the dorsal side of the torso were perceived as coming more from the spine. Further research showed that this bias was not caused by the experimental set-up, the visual system, the subjective location of the stimuli, or other anomalies. The most probable explanation is the existence of two internal reference points: one for the left side of the torso, and one for the right side. When these internal reference points are determined as function of the body side stimulated, the left and right points are 6.2cm apart on average across the ten participants, see Figure 8.

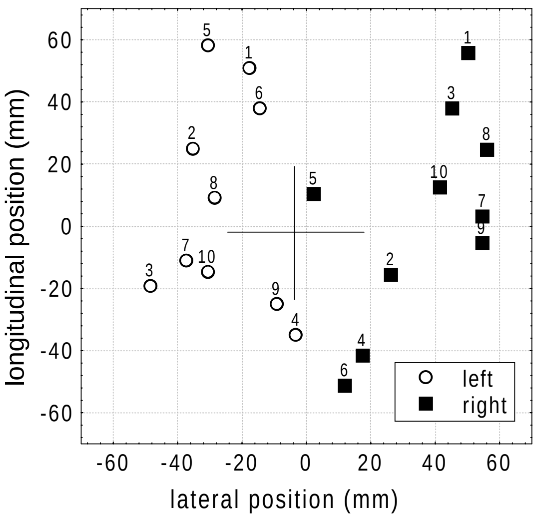


Figure 8: The Internal Reference Points for the ten observers in the tactile direction determination task

The third noteworthy observation is related to the variance of the responses as function of the presented direction. As Figure 9 shows (lower values indicate better performance), scores in the front-sagittal region (–50°—+50° in the graph) are very good with standard deviations between 4° and 8°, and somewhat lower towards the sides.

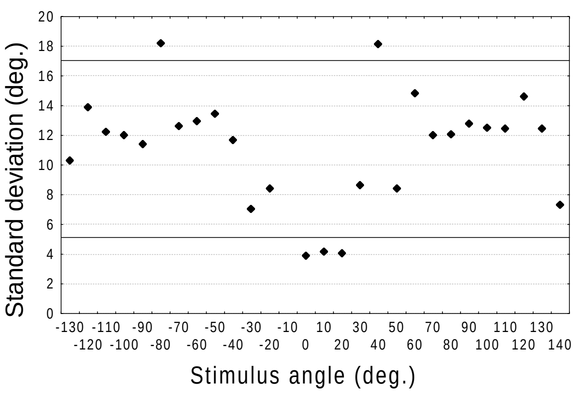


Figure 9: Standard Deviation of the tactile responses as function of the stimulus angle. The horizontal lines summarise the results of the post hoc test; pairs of data points significantly differ when separated by two lines

Other experiments and analysis with the same display are discussed more elaborately elsewhere (Van Erp, 2000). Relevant implications for the application of tactile displays for spatial information are the following:

- observers can perceive a single external tactile point stimulus as an indication of direction,
- although the consistency in the perceived direction varies with body location, performance near the sagittal plane (SD of 4°) is as good as with a comparable visual display,
- direction indication presented by the illusion of apparent location (the percept of one point stimuli located in between two simultaneously presented stimuli) is as good as that of real points,
- small changes in the perceived direction can be evoked by presenting one point stimulus to the frontal side, and one to the dorsal side of the observer.

5. Discussion

Potential beneficial areas of tactile displays in VE systems were presented in Chapter 1. After choosing what information the tactile display must be designed for to present, the relevant perceptual characteristics of the users must be determined. Although there is substantial literature on tactile perception, the available knowledge isn't by far as complete as on visual and auditive perception. Gaps in the required knowledge, e.g. on tactile perception of body loci other than the arms, hands, and fingers, must be filled before applications can be successful. Besides data on fundamental issues such as spatial and temporal resolution, perceptual illusions might be an interesting area in relation to display design. Illusions such as apparent position (which may double the spatial resolution of a display), and apparent motion (which allows to present the percept of a moving stimulus without moving the actuators) offer great opportunities to present information efficiently. Still more illusions are discovered (e.g., Cholewiak & Collins, 1999). After cataloguing all relevant basic knowledge, specific applications must be studied to further optimise information presentation and display use. Another important point, which is not fully addressed in this paper, is the interaction between the sensory modalities, and sensory congruency. An enhanced VE will be multi-modal, but the interaction between the tactile and the other senses is an area, which is only recently being addressed.

When these steps are taken carefully, tactile displays may enhance the experience and effectiveness of the VR.

References

Bach-Y-Rita, P. (1972). *Brain mechanisms in sensory substitution*. New York: Academic.

Bliss, J.C., Katcher, M.H., Rogers, C.H. & Shepard, R.P. (1970). Optical-to-tactile image conversion for the blind. *IEEE Transactions on Man-Machine Systems, MMS-11, 1,* 58-64.

Bolanowski, S.J., Gescheider, G.A., Verrillo, R.T. & Checkosky, C.M. (1988). Four channels mediate the mechanical aspects of touch. *Journal of Acoustical Society of America, 84 (5),* 1680-1694.

Boman, D.K. (1995). International Survey: Virtual-Environment Research. *IEEE, June 1995,* pp 57-65.

Cholewiak, R.W. & Collins, A.A. (1999, in press). The generation of vibrotactile patterns on a linear array: influences of body site, time, and presentation mode. *Perception and Psychophysics*.

Craig, J.C. (1974). Pictorial and abstract pictorial cutaneous displays. In F.A. Geldard (Eds.),

Cutaneous communication systems and devices. Austin, Texas, psychonomic society.

Craig, J.C. (1980). Modes of vibrotactile pattern recognition. *Journal of Experimental Psychology, Human Perception and Performance, 6 (1),* 151-166.

Gilliland, K., Schlegel, R.E. (1994). Tactile Stimulation of the Human Head for Information Display. *Human Factors, 36 (4),* 700-717.

Johnson, K.O. & Phillips, J.R. (1981). Tactile spatial resolution. I. Two point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology*, 6(6), 1177-1191.

Jones, M.B. & Vierck, C.J. (1973). Length discrimination on the skin. *American Journal of Psychology, 86*, 49-60.

Kaczmarek, K.A. & Bach-Y-Rita, P. (1995). Tactile displays. In W. Barfield & T.A. Furness III (Eds.), *Virtual environments and interface design* (pp. 349–414). New York: Oxford University Press.

Kandel, E.R., Schwartz, J.H. & Jessel, T.M. (1991). *Principles of Neural Science*. New York: Elsevier Science Publishing Co.

Linvill, J.G. & Bliss, J.C. (1966). A direct translation reading aid for the blind. *Proceedings IEEE, 54,* 40-51.

Loomis, J.M. (1974). Tactile letter recognition under different modes of stimulus representation. *Perception and Psychophysics, 16,* 401-408.

Loomis, J.M. & Collins, C.C. (1978). Sensitivity to shifts of a point stimulus: An instance of tactile hyper acuity. *Perception and Psychophysics, 24,* 487-492.

O'Modrain, M.S. & Gillispie, B. (1997). The moose: a haptic user interface for blind persons. Internet.

Roelofs, C.O. (1959). Considerations on the visual egocentre. *Acta Psychologica, 16,* 226-234.

Rupert, A.H., Guedry, F.E. & Reschke, M.F. (1993). The use of a tactile interface to convey position and motion perceptions. *AGARD meeting on Virtual interfaces: reserach and applications*, October 1993.

Snippe, H.P. (1991). *Human perception of spatial and temporal luminance structure*. PhD thesis, Utrecht University, The Netherlands.

Starkiewicz, W., Kuprianowicz, W. & Petruczenko, F. (1971). 60-channel elektroftalm with CdSO4mphoto-resistors and forehead tactile elements. In Sterling et al. (Eds.), *Visual prothesis: the interdisciplinary dialogue*. New York: Academic press, 295-299.

Valbo, Å.B. & Johansson, R.S. (1978). The tactile sensory innervation of the glabrous skin of the human hand. In Gordon (Ed.), *Active touch: the mechanisms of recognition of objects by manipulation. A multi-disciplinary approach* (pp. 29-54). Oxford: Pergamon press.

Van Erp, J.B.F. (2000, in press). *Accuracy of and bias in 2D direction perception of tactile stimuli presented to the torso* (Report TM-00-B0?). Soesterberg, The Netherlands: TNO Human Factors.

Van Erp, J.B.F. & Dobbelsteen, J.J. van (1998). *On the design of tactile displays* (Report TM-98-B012). Soesterberg, The Netherlands: TNO Human Factors.

Van Erp, J.B.F. & Vogels, I.M.L.C. (1998). *Vibrotactile perception: a literature review* (Report TM-98-B011). Soesterberg, The Netherlands: TNO Human Factors.

Van Erp, J.B.F. & Werkhoven, P.J. (1999). *Spatial characteristics of virbo-tactile perception on the torso* (Report TM-99-B007). Soesterberg, The Netherlands: TNO Human Factors.

Vierck, C.J. & Jones, M.B. (1969). Size discrimination on the skin. *Science*, 163, 488-489.

Vierodt, K.H. (1870). Abhängigkeit der Ausbildung des Raumsinnes der Haut von der Beweglichkeit der Körpertheile. (Dependence of the development of the skin's spatial sense on the flexibility of parts of the body). *Zeitschrift für Biologie*, 6, 53-72.

Weber, E.H. (1834). De pulsu, resorptione, auditu et tactu. In H.E. Ross & D.J. Murray (Eds.), *E.H. Weber, on the tactile senses*. Hove, UK: Taylor & Francis.

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body-part, sex and laterality. In D.R. Kenshalo (Ed.), *The Skin Senses* (pp 195-218). Springfield: C.C. Thomas.

Werkhoven, P.J. & Van Erp, J.B.F. (1998). *Perception of vibro-tactile asynchronies* (Report TM-1998-B013). Soesterberg, The Netherlands: TNO Human Factors.

Ziegler, R. (1996). Haptic Displays — How can we feel Virtual Environments? — Imaging Sciences and Display Technology, *SPIE Proceedings vol. 2949*, pp 221-232.

This page has been deliberately left blank

Page intentionnellement blanche

Virtual Cockpit Simulation for Pilot Training

Dipl-Ing. Kai-Uwe Dörr, Dipl.Inform. Jens Schiefele & Prof. Dr.-Ing. W. Kubbat[1]
Institute for Flight Mechanics and Control
Technical University Darmstadt
Petersenstraße 30
64287 Darmstadt
Germany

Summary
For some of today's simulations very expensive, heavy, and large equipment is needed. Examples are driving, shipping, and flight simulators with huge and expensive visual and motion systems.
In order to reduce cost, immersive 'Virtual Simulation' becomes very attractive. Head Mounted Displays (HMD) or CAVEs (Computer Animated Virtual Environments), Datagloves, and cheap 'SeatingBucks' are used to generate a stereoscopic virtual environment (VE) for the trainee.
IVS enhances training quality and quantity for classroom-teaching and Computer Based Training (CBT). It allows to visualize and animate teaching-material in a more natural stereoscopic environment. Data of before unseen complexity can be revealed and complex models easily visualized. For the first time, the trainee himself can interact with a Data-Glove in the environment and collect cockpit experience long before his maiden flight. CAVEs and Immersive Projection Screens enable "group training" to collect personal and shared experience while further enhancing training quality.
With increasing maturity of VR-gear IVS will allow to generate new training metaphors for immersive flight simulation. This might include the enhancement or partial replacement of conventional flight simulators by IVS.

Introduction
High fidelity pilot training simulators are designed as training tools for one specific aircraft type. They demand authentic instrumentation and system layout for the simulated aircraft type including huge outside vision systems and cumbersome motion systems.[1] Because of these reasons, traditional simulators are very expensive, inflexible, and difficult to reconfigure. The high cost factor in buying and maintaining them causes air carriers to purchase either just a single simulator for every aircraft type they own or to buy expensive training hours from other companies.[1]

Virtual Simulation
To overcome some of the problems in the field of pilot training the Air Force Institute of Technology developed a Virtual Cockpit (VC) for fighter pilot training.[2] Pilots are immersed in a stereoscopic VE, wearing a HMD and a pointing device to interact with the virtual cockpit devices.[3] A VC can be easily reconfigured by simply switching the cockpit model database and the attached flight mechanics.[4] At the Institute for Flight Mechanics and Control, Darmstadt University of Technology this concept was extended to be suitable for an Airbus A340 Cockpit-IVS using hi-resolution HMD, "Seating Buck", cyberglove, and stereoscopic projection screens for a natural interaction metaphor.[5] As pilot outputs for principle navigation and Instrument Flight Rule (IFR) testing a virtual Primary Flight Display (PFD), a virtual Navigation Display (ND), and a virtual civil Head-Up Display (HUD) are available. In addition, a simplified outside visual is rendered to the pilot. These displays are sufficient to run principle Instrument Flight Rule (IFR) tests with the virtual cockpit.
The problem of lacking force feedback in IVS was significantly reduced by developing a "Seating Buck".[6] Only side-stick, pedals, flap-lever, and thrust-lever are physically available. All other buttons, dials, and switches are simulated by simple plastic panels. In a test series the concept and implementation proved to reduce interaction time significantly.[1,6]
Other examples for operator training using Virtual Training methods are Astronaut training to repair the Hubble Space Telescope[7], submarine outlook training to practice maneuvering in a harbor[8], support pilots classroom education[9], or caterpillar training. Instead of HMDs very often CAVEs[10] and BOOMs[11] are used to avoid heavy intrusive head gear and limited Field of Views (FOV).

Human Machine Interface in Virtual Simulation
Transfer of training from virtual into real space still has to be proven for pilot training. For simple Cola can sorting in a CAVE transfer of training from virtual into real space was shown.[12,13] Also, people trained in VR have a better orientation in buildings than map trained persons[14]. Therefore, it can be assumed that training in virtual environments might be useful to train trainees at different requirement levels.
Training quality limiting factors due to today's hardware equipment such as Field of View (FOV)[15,16,17,21], tracker

[1] For contact with authors: doerr@fmr.tu-darmstadt.de; schief@fmr.tu-darmstadt.de; kubbat@fmr.tu-darmstadt.de; tel. +49 6151 16 2890, fax +49 6151 16 5434

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

latency[18], presence[19], and missing force feedback[1,6] were investigated in principal. Only few research exists determining the limitations on training caused by the complete VR-Human Machine Interface (HMI) design and hardware.[20]

VR-HMI research has already been conducted concerning force feedback, HMD FOV, and HMD resolution for Cockpit-IVS.[1,6,21]

A good general overview describing VR-HMI research is presented in [22].

Conventional Computer-Based Training

Computer Based Training and Procedure Training (PT) use PCs with a 2D image, sound, a mouse, and a keyboard for interaction. A trainee sits in-front of a PC screen and interacts by clicking with the mouse. CBT is split into different chapters such as radio navigation, flight planing, flight performance, electronics, instrumentation, and engines. Further enhanced systems allow partial simulation of functionality. For each individual aircraft type a different program is available. Each individual chapter is split into different learning units:

- Overview
- Components and Control
- System Operation
- Abnormal Operation
- Summary
- Mastery Test

In different learning units the trainee gets a multimedia presentation of the learning material. After the introduction, the trainee can interact with the system by clicking with the mouse on interaction devices. In the Mastery Test multiple choice questions have to be answered and tasks performed. The trainee can practice all units on his own personal learning pace. The test can also be individually repeated.

Such a training metaphor helps to support individual training. “Fast learner” are not frustrated by a low pace and “slow learner” are not overrun.

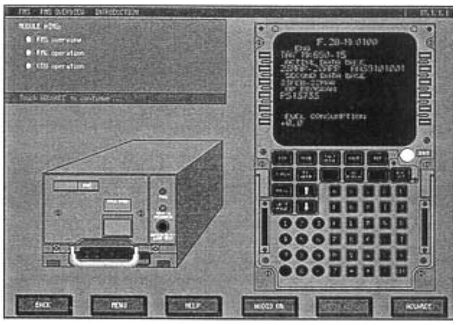


Figure 1: Flight management computer training with CBT

The trainee does not have any immersive experience towards the real geometry and functionality of the cockpit. The position of interaction devices in real 3D space is unknown to him. Familiarization in 3D space can not be realized with today’s CBT systems.

CBT with a VR Training Environment

In order to enhance CBT, a 3D Virtual Cockpit model is generated. All interaction devices such as side stick, pedals, thrust-lever, knobs, buttons, and dials are modeled as 3D geometry. All other parts and surfaces are formed by simple textured geometry.[5] This 3D model is rendered to a pilot wearing a tracked high-resolution, large field of view (FOV), stereoscopic HMD.

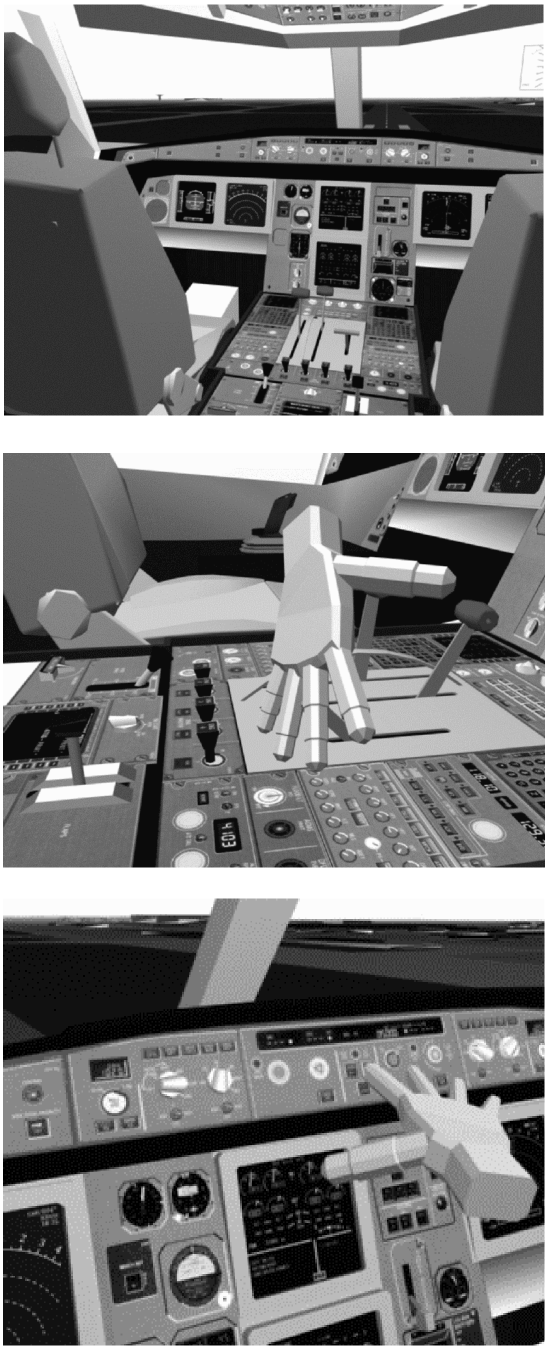


Figure 2: IVS Cockpit based on modeled geometry and textures

For interaction the pilot wears a tracked data-glove recognizing hand position, orientation, and finger bending. The trainee can virtually interact with all cockpit devices, dials, and buttons. The system response on the input can be visualized.

The same image is also rendered to a large stereoscopic projection screen (Shutter classes) enabling observers to watch the trainee and his interaction. This allows later discussion on the trainees performance.

Figure 3: Demonstration room with projection screen (three shutter signals)

The same concepts and structures known from ordinary CBT are applied. The only difference is that the trainee is immersed into the scene allowing him to interact naturally with his environment. Learning aids and eye catchers such as symbols, markers, and any desired virtual information can be visualized within the 3D virtual cockpit as well. For instance, after toggling the gear lever the virtual gear unit is displayed and the actuator changes visualized. Therefore, beyond the VR simulation of an ordinary cockpit, virtual information can be incorporated and used as didactical metaphor for pilot training.

Therefore two different Training methods are feasible with this environment. In the so called Class Room Training, collective and cooperative learning in front of a single projection screen enables trainees to work and learn together in the same environment.

As a second training method, a single VR-Pilot training environment was developed. Therefore the trainee wears a HMD and a Data Glove. Both devices were tracked with a tracking system. So a naturally interaction with the virtual scene is possible. The trainee is guided trough the different lessons depending on his interaction with the cockpit The trainee itself define which lesson he wants to work through . Additionally it is possible for the trainee to fly with this virtual cockpit because of its full functionality. Both Methods uses the same Training lessons with minimum changes in interaction possibilities. Example lectures were realized for both training methods and will be described later on.

Classroom Training

The didactical methods for training vary depending on the airlines and the training facilities. Training is often based on the conventional concept of "frontal teaching". Teachers give lectures with varying didactical materials. Dependent on the training facility this can be simple transparencies, video-tapes, sketches, boards, and small mockups. After each lecture the trainee has the possibility to re-read the taught lecture from printed material. At the end of each training chapters pilots must pass a written multiple choice test.

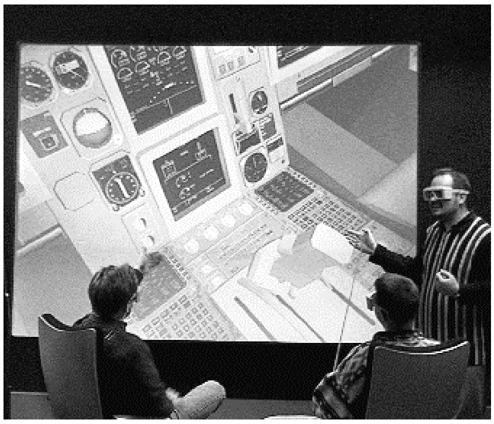


Figure 4: Stereoscopic projection screen for classroom training

Hence, the understanding of complex aircraft systems strongly depends on the imagination of a trainee and the teaching skills of a teacher. In order to enhance teaching quality stereoscopic projection screens can be used to visualize complex aircraft systems and technical dependencies in a natural way. A teacher can fly through a model, hide obstructing parts, or animate complex functionality's. The trainee himself becomes part of the scene. In after lecture sessions the trainee can interact with the system and its functionality. In front of an immersive screen the trainee becomes part of the scene and experiences the learn material more naturally. Stereoscopic vision with depth perception enables new pilots to easily asses complex 3D structures, aircraft positions, etc. The "hands-on experience" helps to deepen the understanding and motivate the trainee to explore deeper into the learning material. Group experience and group training can be enhanced with an stereoscopic projection system. This might accelerate memorization and pushes the later needed ability of cooperative cockpit work through group experience.

VR Pilot Training

For today's simulator training huge and expensive simulators are needed. Each training hour costs up to $5,000 and existing training facilities are currently at the limits of their capacities. Additional to the later on described training lessons, the system is also fit for use, to train some real flight tasks. Therefore, the Virtual Cockpit (VC) based on the above described technique (HMD plus Stereoscopic Projection Screen) can be used. In addition to the CBT approach, a simplified outside visual is added to generate an immersive flight simulation. The viewing distance has to be reduced to approximately 20km in order to ensure sufficient rendering performance (15–20Hz). A virtual Primary Flight Display (PFD), Navigation Display (ND), and a virtual stereoscopic Head Up Display (HUD) are used in a first approach.[21]

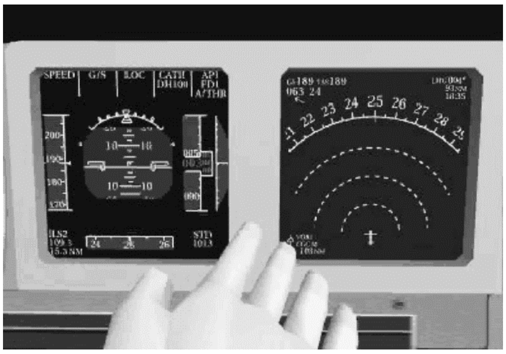


Figure 5: Primary flight display and navigation display

These virtual displays show basic information necessary to perform a controlled flight and allow basic performance analysis with the system.

Aircraft System Lesson

Aircraft system knowledge is one of the main training parts during theoretical pilot education. Simple system diagrams are used to give the trainees an overview of the whole system. Relations between aircraft subsystems and how these systems work together will be explained in the same way. This is not a very intuitive way of learning. The best way to learn is to visualize them. The visual channel is the most significant way to comprehend information.

The main advantage of VR-systems is the possibility to display trainees a 3D geometry of an object and a simulation of the real behavior. As an example the behavior of gear, flaps, and rudder on an input from the pilot is shown. Therefore, a complete aircraft model is shown to the trainee (Figure 6). The model shows the reaction of the aircraft and it is possible for the trainee to zoom in different subsystems.

Figure 6: Aircraft outside view above the pedestal

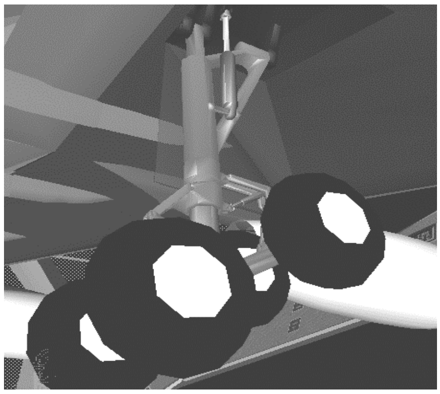


Figure 7: Gear view

In this example (Figure 7) the trainee can observe the kinematics of the gear. He can imagine the 3D behavior of the actuators during gear up and down procedure. So, in case of a system failure, he can imagine what is happen and where the errors can be. The system knowledge increases because of the 3 dimensional form of presentation.

Engine Lesson

During pilot education engines are visualized by explosion sketches or vertical cuts through an engine. This creates a complex visualization of the parts. For instance, explaining the turbine turn rate at N1 and N2 is rather difficult. Either the graphical representation is showing too much or too view detail, forcing the instructor to switch between several images.

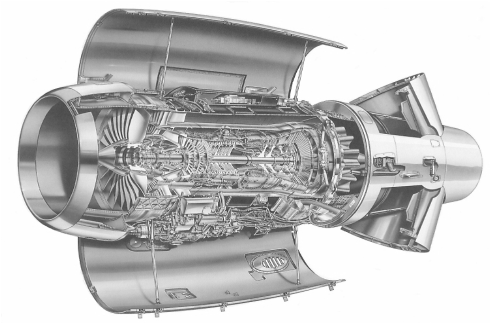


Figure 8: Conventional visualization with a cut through an engine

Therefore, a lecture was developed that allows to dismantle the engine from a full blown representation down to the necessary key elements such as fan, turbine stators, turbine shaft, and combustion chamber. Trainees observe the animated engine and can position themselves on arbitrary positions.

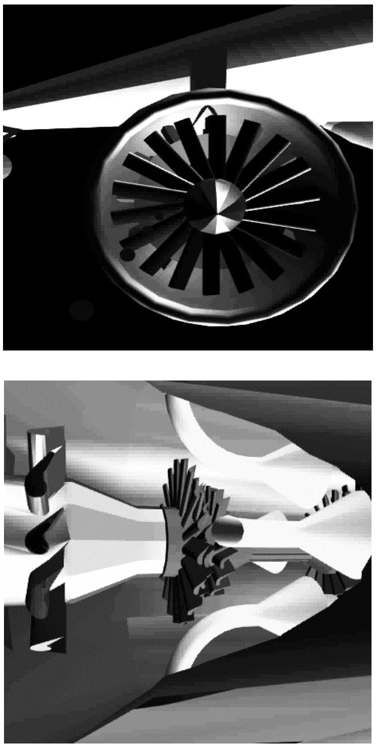


Figure 9: 3D engine inside view

Alternatively the they can follow a prerecorded flight path through the scene. During the flight they can arbitrary change the viewing direction. They can always stop and move towards any object to get a closer look. To increase realism and a feeling for object sizes, a complete aircraft is rendered as well.

Stereoscopic vision enables trainees to be immersed into the environment generating a closer and better impression of the turbine. It enables the trainees to achieve a feeling for real part and turbine sizes. With ordinary paper sketches this is impossible. As an enhancement, observers can be re-scaled to small sizes allowing to closely observe small turbine parts and their functionality.

Based on the stereoscopic vision, instrument locations and attached functionality can be memorized by generating a mental map of the cockpit.

Force Feedback/Vision

It was determined that lacking force feedback in pure IVS is a major usability limitation.[6] Therefore, some devices are physically available such as sidestick, pedals, and thrust-lever. All others are replaced by simple plastic panels to generate force feedback to the pilot (Seating Buck).[1] The Seating Buck can be easily reconfigured to simulate arbitrary cockpit configurations.

With a Seating Buck force feedback device the interaction time is reduced significantly providing a more natural haptical feedback to the pilot.[6]

For the success of a VR CBT enhancement a large FOV of more than 80° is needed.[21] Above a 60° FOV pilots can assess all visible information and geometry in the cockpit. Above a 80° FOV also orientation and cross-viewing among two pilots simulated in the same IVS is feasible.[21]

Figure 10: Stereoscopic projection screen rendering scene visible to the pilot

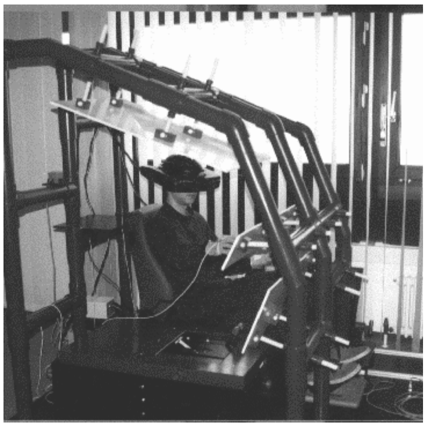


Figure 11: Seating buck to simulate force feedback

Usability

VR CBT and PT system can help to reduce education cost by reducing expensive simulator hours for familiarization and principle interaction training. Otherwise, it can serve as an extension to already proven CBT training concepts. The VR technology and projection screen technology is mature enough to fulfil these tasks. HMD with the requested FOV, force feedback devices, and computers with sufficient graphics power exist. However, real verification of training transfer has to be further investigated in the future.

Flight Simulation

All ordinary software simulation modules known from conventional flight simulation such as physical input devices, virtual input devices, flight mechanics, traffic, and rendering run in a distributed environment on different high end graphics work stations. As simulation module an Airbus A300 flight mechanics, ground collision , weather, and sound modules are available. All the modules are taken from a conventional flight simulator available at the Institute for Flight Mechanics and Control.[23] It can be also used for comparing the VCS with a real flight simulation.

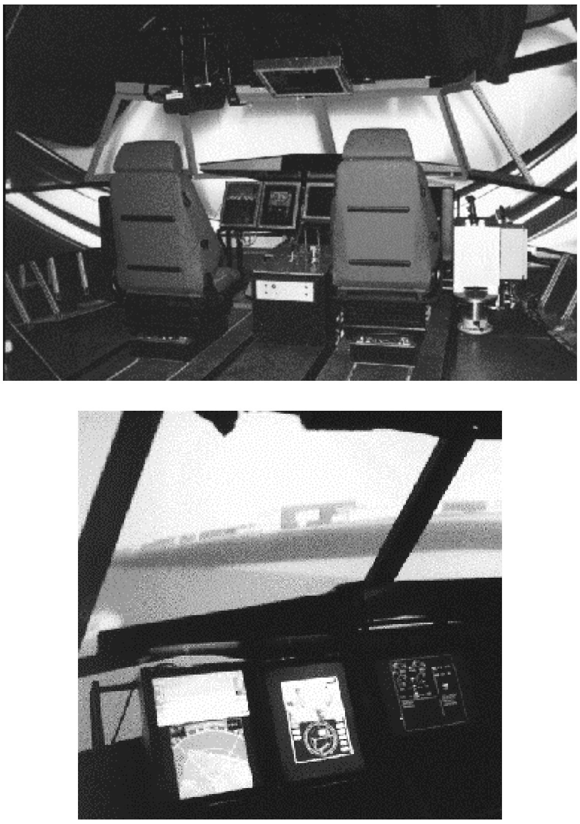


Figure 12: Conventional flight simulator mock-up at the FMRT

Motion Base

In addition, to the current approach of a fixed based “Seating Buck”, the entire system can be mounted to a motion base. This would increase the level of immersiveness by aircraft motion. To simulate a civil aircraft the performance of small two-seater (300kg) motion bases would be sufficient. The increase in realism and immersion is untested and need to be further evaluated.

Tracker and System Lag/HMD Limitations

One of the key limitations to VC is today’s tracker latency. The entire VCS has currently a latency of about 100ms.[24] From tests it was deducted that 150ms are sufficient for orientation tasks within the cockpit.[24] Maximal lag for flying a VCS should be well below 80ms.[25]

Another limitation is the currently used HMD with a FOV of maximal 56°. As stated above, this HMD reduces the FOV in a way that disables flying and orientation tasks within the cockpit. However, this is not a general concept limitation because HMD vendors already sell 120° equipment.

HMD resolution is very critical for the usability of VCS. Research results indicate that with a hires HMD of 1280x1024 pixels and cockpit displays (PFD, ND, and HUD) of 8 inch rendered at a distance of approximately 85cm (standard pilot-display distance) offers sufficient resolution.[21] Therefore, resolution with modern HMD is no more limitation to Cockpit-IVS.

Usability

The used VR equipment is critical for the success of VCS. In principal most system components such as force feedback generation, HMD FOV and resolution are proven to be usable. Tracker lack has to be reduced significantly. It is untested which influence the combination of VR equipment has on the overall performance. The usability of VCS to enhance simulator training is unproven.

Even after optimization of all VR equipment it seems unfeasible to completely replace flight simulators by VCS. In a first step the understanding of the HMI presented by intrusive, heavy, inconvenient Virtual Reality gear has to be further investigated. Also, a complete Virtual Reality simulation theory is missing. On the first impression large projection screens have less negative impact on the HMI.

Future Work

The Institute for Flight Mechanics and Control will further investigate the usability of VR-CBT and VCS. The research will be focused on the human machine interface generated through virtual simulations. The goal is to prove the usability (especially of CBT) for real usage in today’s flight training.

It is assumed that the introduction of large stereoscopic projection screens into today’s pilot training will be a natural step. The equipment is already usable for classroom training applications and CBT.

Available Equipment

At the Institute for Flight Mechanics and Control a variety of equipment can be used. A front projection system (Ampro Projector) with a curved screen and four shutter emitters is installed. For the system seven shutter glasses (Christal Eye) are available. For IVS a nVision Datavisor 10× (1280×1024), Kaiser ViSim500, Polhemus Fastrack, two 18 sensor Cyber-Gloves, and a triple Pipe Onyx (IR) can be used.

Acknowledgements

We would like to thank the Fraunhofer Institute for Computer Graphics (IGD) and their software distributor VRcom in Darmstadt for supporting us with their Virtual Reality software package “Virtual Design II”.[26]

References

[1] J. Schiefele et al., “Virtual Cockpit Simulation with Force Feedback for Prototyping and Training”, Society for Image Generation, Scottsdale, JS1-JS9, 1998

[2] McCarthy et al., “A Virtual Cockpit for a Distributed Interactive Simulation”, IEEE Computer Graphics & Applications, pp. 49-54, 1994

[3] M. Stytz et al., “The photo realistic Virtual Cockpit”, SPIE/SCS Joint 1996, Government & Aerospace Simulation Conference, pp. 71-76, New Orleans, 1996

[4] M. Stytz et al., “Issues in the development of rapidly reconfigurable immersive human-operated systems for distributed virtual environments”, SPIE vol. 3085, pp. 162-172, Orlando, 1997

[5] J. Schiefele et al., “IFR Cockpit Simulation in a Distributed Virtual Environment”, SPIE vol. 3067, Orlando, 1998

[6] J. Schiefele et al., “Simple Force Feedback for Virtual Cockpit Environments”, SPIE vol. 3067, Orlando, 1998

[7] R.B. Loftin, “Virtual Reality for Aerospace Training”, VR Systems Magazine, 1996

[8] DoD, “Virtual Environment for Submarine Ship Handling”, Res. and Eval Program 6.2, Virtual Environment and Training, DoD, 1997

[9] J. Schiefele et al., “Stereoscopic Projection Screens and Virtual Cockpit Simulation for Pilot Training”, Immersive Projection Screen Technology (IPT), pp.211-222, Stuttgart, 1999

[10] Cruz-Neira et al., “Surround Screen Projection based VR. The Design and Implementation of the CAVE”, SIGGRAPH, pp. 135-142, 1993

[11] M. Bolas, “Human Factors in the Design of an Immersive Display”, IEEE Computer Graphics & Applications, pp. 55-59, 1994

[12] Kozak, “Transfer of Training from Virtual Reality”, Ergon 36, 1993

[13] R. Kenyan, M. Afenya, “Training in Virtual and Real Environments”, Technical Publication, Univ. of Illinois, Chicago, 1997

[14] B. Knerr et al., “Training in Virtual Reality: Human Performance, Training Transfer, and Side Effects”, Society for Image Generation, Scottsdale, 1996

[15] E. Kasper, “Effects of in flight field of view restriction on rotorcraft pilot head movement”, SPIE, Vol. 3058, Orlando, 1997

[16] Z. Szoboszlay et al., “Predicting Usable FOV Limits for Future Rotorcraft Helmet Mounted Displays”, DERA, UK, 1997

[17] K. Arthur, “Effects of FOV on Task Performance with Head Mounted Displays”, University of North Carolina, 1997

[18] S. Rogers et al., “Effects of System lag on head tracked cursor control”, SPIE, Vol. 3058, Orlando, 1997

[19] B. Witmer, “Measuring Presence in Virtual Environments: A presence questionnaire”, Presence, Vol. 7, MIT-Press, pp. 225-240, 1998

[20] M. Snow, “Charting Presence in Virtual Environments and it’s Effects on the Performance”, Dissertation, Virginia Polytec Institute and State University, 1996

[21] J. Schiefele et al., “Evaluation of required HMD resolution and FOV for Virtual Cockpit Simulation”, SPIE, Vol. 3689, Orlando 1999

[22] K. Stanney et al., “Human Factor Issues in Virtual Environments”, Presence, Vol. 4, MIT-Press, pp. 327-351, 1998

[23] O. Albert et al., “Das Forschungscockpit der TU-Darmstadt, ein Werkzeug zur Untersuchung neuer Cockpitkonzepte”, DGLR Fachausschußsitzung für Luft- und Raumfahrt, 1998

[24] A. Eichler, “Untersuchung des Einflusses der Tracking Totzeit einer virtuellen Simulation auf die Leistungsfähigkeit der Benutzer im Hinblick auf künftige VR-Cockpitsimulationen”, Study Thesis, TU-Darmstadt, 1999

[25] F. Brooks, “What is real about Virtual Reality?”, IEEE-VR, 1999

[26] P. Astheimer et al., “Virtual Design II: An advanced VR System for Industrial Applications”, Virtual Reality World, Stuttgart, pp. 337-363, 1995

This page has been deliberately left blank

Page intentionnellement blanche

UAV Operations using Virtual Environments

Jan B.F. van Erp & Leo van Breda[1]
TNO Human Factors
Kampweg 5
3769 DE Soesterberg
The Netherlands

Summary

In virtual environments (VE), the limited field of view, the lack of information on viewing direction, and possible transmission delays may be considered as potential problems in developing and maintaining a good sense of situation awareness. Enabling unmanned air vehicle (UAV) operators to use high quality (proprioceptive) information on (changes in) viewing direction by introducing a head-slaved camera system with head-slaved display (HMD) may improve situation awareness, compared to using a joystick and a fixed monitor. However, HMDs may degrade comfort and the dynamics of head movements. Furthermore, time delays and zoomed-in images induce a non-steady presentation of the environment, and may impede adequate mapping of spatial information. This paper reports an exploratory study into the applicability of a head-slaved camera system in unmanned platform applications. To overcome the possible drawbacks of HMDs, we compared an HMD with a head-slaved dome projection in a simulator experiment. To overcome the possible drawbacks of transmission delay, we introduced a new method to compensate for the spatial distortions. This technique, called delay-handling, preserves the correct spatial relation between the viewing direction of the camera and operator by presenting incoming images in the camera viewing direction, and not in the actual viewing direction of the operator.

The experimental results showed that delay-handling is successful in supporting the perception of correct spatial relations, i.e., it improves situation awareness. No differences in task performance were found between the actual HMD and the dome projection.

Introduction

In operating a Maritime Unmanned Aerial Vehicle (MUAV) the flow of information is very poor as compared to real flying. If a human operator was physically present at the remote site and performs manipulations directly, he would receive a variety of information on the result of his manipulations, such as visual, auditory, tactile, and force feedback. However, when the human is physically separated from the task space, the feedback of the control actions has to be artificially transmitted back to him.

The man-machine interface determines the extent to which the operator can sense the remote environment and consequently control the platform. Thus, the display and controls in the operator environment should be designed in such a way that the operator receives task specific information and sufficient feedback. The images provided by an on board camera is the main source of information on the outside world for MUAV operators. Because of the inherent characteristics of a camera-monitor system, and the restricted data link between the remote site and the operator, these images are of degraded quality, which may affect steering and control performance and the operator's situation awareness (SA).

Image degradation may come in different forms, e.g. a reduced field of view, a zoomed-in image, decreased information about the camera viewpoint and viewing direction, a time delay between the control input and the consequent feedback, and reduced spatial and temporal resolution. It is plausible that the degradation of some aspects of the feedback is more detrimental for operator performance or the sense of SA than others; some information may be redundant or of only secondary value. In order to identify the limitations that may become critical for the sense of SA when the operator manually controls MUAV and/or camera movements we first reflect on the concept of SA. Next, regarding MUAV operators, the main issues that affect SA will be discussed. Finally, we establish which principles of interface design may support the operator in developing a good sense of SA.

In teleoperation, situation awareness may be defined as the operator's ability to perceive, comprehend, and predict the spatial layout of the elements in the environment. SA is not a static phenomenon, but is composed of a variety of changing facts, interpretations and predictions in the context of task requirements. Although operator performance undoubtedly depends on SA, their exact relationship is not clear. Actually, there is still disagreement among researchers as to just what constitutes SA. However, the elements of SA are well known and include such familiar human functions as perception, information processing, decision-making, memory, learning, and action-taking, performed within a dynamic set of environmental circumstances and conditions.

SA is important in a wide variety of environments. Acquiring and maintaining SA becomes increasingly difficult as the complexity and dynamics of the environment increase. Under some circumstances, many decisions are required within a fairly narrow time span, and task performance requires an up-to-date analysis of the environment. Because the state of the environment is

[1] For contact with authors: vanerp@tm.tno.nl and vanbreda@tm.tno.nl

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

constantly changing (often in complex ways) a major portion of the operator's job becomes that of obtaining and maintaining good SA.

Barfield, Rosenberg and Furness (1995) describe the main components of situation awareness: spatial, status, and overall situation awareness. Spatial or navigational awareness deals with the three-dimensional geometry of the environment and refers to the operator's mental model of the vehicle's position. What is my position and how does this relate to the position of other objects? The state of the platform, e.g. the amount of remaining fuel, the position of the flaps, is represented in the status component of awareness. The combination of spatial and status awareness enables an overall awareness of the total flight environment.

Endsley (1995) gives a more elaborated model of SA with three components. Level one in this model refers to the perception of the elements in the environment and their relationship to other points of reference (i.e. internal model). At this level, relevant characteristics (colour, size, speed and location) and the dynamics of the objects in the environment are represented. This aspect is similar to what Barfield et al. (1995) termed spatial awareness. Level two of SA goes beyond simply being aware of the elements that are present, and includes an understanding of the significance of the elements. Based on level one knowledge, the operator forms a holistic picture of the environment, comprehending the significance of objects and events. Thus, the integration of various level one data elements at level two of SA is crucial for the comprehension of the situation. Level two of SA can be highly spatial in an operating context. The relevance of different objects for the operator's action planning will depend on their location and speed. Finally, the ability to project the future actions of the elements in the environment forms the third and highest level of SA. For example, in traffic, knowledge of the status and dynamics, and the comprehension of the situation, allows a driver to predict the future actions of other drivers in order to prevent collisions.

Another aspect of SA should be mentioned at this point. Although SA has been defined as a person's knowledge of the environment at a given point in time, it is highly temporal in nature. That is, some aspects, like the knowledge about the dynamics of the environment and path prediction, are acquirable only over time. Smolensky (1993) discusses the work of Stein, who showed that controller's eye fixation locations, which had varied widely in the initial 10 to 15 minutes of an air traffic simulation, decreased significantly beyond that point in time. Anecdotally, Stein's subjects reported that the initial 10 to 15 minutes of a controllers shift is the period of time during which he acquires the 'big picture', or, SA. Another temporal aspect of SA relates to the variations in relevance of elements across time. Some elements are not of equal importance at all times, although they should not fall out of consideration completely. At least some SA on all elements is needed. SA, therefore, is based on far more than simply the information perceived about the environment. It is related to a model of human information processing in which attention and long-term memory enable comprehending the meaning of information in an integrated form. Memory does not only serve to direct attention effectively, but also serves to interpret the information that is perceived and to develop accurate projections of future events.

SA in teleoperation

In teleoperation, an intervening system senses, mediates, and presents information to the human operator. In this process, a loss of information can occur, which may be relevant to all three levels of SA.

At the lowest level, the system may fail to present certain information that is important for SA in the assigned task. First, systems may only present information of one modality (e.g. only visual information), based on technological limitations and the designer's understanding of what is required. Second, the information that is presented may lack important cues; e.g. no stereoscopic depth cues when a single camera is used.

Another major issue in teleoperation is the transmission speed and capacity. Intervening communication systems like satellites reduce transmission speed, resulting in delayed feedback to the operator about his manipulations.

For level two SA, the information displayed by the system must be integrated, and related to a mental model to obtain a holistic picture, and to determine which cues are actually relevant to the established goals. When no model exists at all, level two SA must be developed in memory. The absence of sufficient level one SA, the inability to develop a sufficient mental model or the inability to properly integrate or comprehend the meaning of presented data, can lead to inaccurate or incomplete level two SA. This may be caused by incomplete or inaccurate presentation of data to the human operator, or by a mismatch between information presentation and perceptual, attentional, and working memory characteristics of the operator.

Finally, level three SA may be lacking or incorrect. Even if the mental model is sufficient for level two SA, and the actual situation is clearly understood, it may be difficult to accurately project future dynamics. Lack of highly developed mental model and attention and memory limitations may account for this. Furthermore, some people are simply not good at mental simulation.

Regarding the control of unmanned platforms, loss of SA is already present at level one of SA, causing degraded sense of SA on level two and three as well. The inability to assess basic properties as position, direction and speed also hampers the operator in developing a correct mental model (level two), and in making adequate predictions about future states of the objects (level three). Part of the problems are probably related to

the poor information flow specific in MUAV applications, due to the following reasons:
A small field of view. A limited field of view suppresses the use of peripheral visual information. The peripheral area of the retina differs anatomically and functionally from the foveal area (Schneider, 1969; Trevarthen, 1968), and is used to generate our sense of spatial orientation (Ungerleider & Mishkin, 1982; Jeannerod, 1997). For example, a human operator's performance in a disturbance nulling task with only a central field of view display can be dramatically improved if the field of view is expanded to cover the peripheral retina (Kenyon & Kneller, 1992).
Furthermore, a small field of view requires a higher degree of integration of spatial information to build up a representation of the spatial environment. That is, rather than having a large field of spatial information in which several objects (and terrain features) are localised, a smaller field of view affords less spatial information at any instant, which forces operators to integrate these small 'pieces' of spatial information in time. The results of a search and replace experiment using an HMD (Venturino & Kunze, 1989) indicated that the field size affects one's ability to acquire spatial information. However, an important observation in this experiment was also that once the spatial information has been mapped into spatial memory, humans could use that information independently of the size of their 'window' to the world. This phenomenon is also found by Thompson (1983), who asked subjects to walk with closed eyes to previously viewed targets, and Tyrell et al. (1993) who asked visually occluded subjects to position a point of light at the location of a previously viewed target.
A zoomed-in image. Often, the small field of view is combined with a zoomed-in camera image. The zoom-factor of the camera disturbs the normal relation between rotational speed of the camera and translational flow in the camera image. For example, Van Erp, Korteling and Kappé (1995) found that operators largely overestimate camera rotations when viewing a zoomed-in camera image.
Few points of reference at sea. The lack of reference points at sea may hinder the operator in developing a good model of the position of objects in the remote environment and their relations.
Low update rate. Update rates lower than 4 Hz limit the perception of the direction and speed of objects, platform and camera.
Transmission delays. Transmission delays will mainly lead to degraded performance of the operator when manually controlling the camera. Eventually, the operator will develop a go-and-wait strategy, which will hamper developing a sense of SA.
Degraded information on (changes in) the viewing direction. Controlling the viewing direction of the camera by means of a joystick while the images are presented on a stationary monitor, withhold the operator of proprioceptive feedback on viewing direction. Normally this information is provided by muscle spindles of neck and eyes, and therefore allows automatic mapping of visual information on a mental model. Since the viewing direction can not be directly deduced from the camera images, it is usually presented via additional indicators. However, this information requires the operator to perform some kind of cognitive processing in order to build a mental model, and it is not intuitive and therefore slow.

In previous research, it was shown that introducing high quality synthetic visual information can partly cancel out problems regarding the zoomed-in camera image, the lack of reference points, the low update rate and the transmission delay, which all have an important camera control component (Van Erp, Kappé & Korteling, 1996).
Field size and information on viewing direction may be considered as the most important factors related to SA in unmanned platform applications. Moreover, both factors probably interact strongly. Although spatial information can be used effectively regardless of the size of the 'window' to the world once it is stored in spatial memory; the lack of information about the viewing direction of the camera hinders the building of a mental representation, and the integration of new information.

Head-slaved camera control

A possibility to convey high quality information about camera viewing direction is the use of a head-slaved camera system. When the viewing direction of the camera is coupled to the viewing direction of the operator, proprioceptive information is available, which can be interpreted automatically. Automatic processing tends to be fast, autonomous, effortless, and unavailable to conscious awareness in that it can occur without attention. It is hypothesised that system designs that support automatic processing of information directly benefit performance.

Applying a head-slaved camera system also requires a head coupled image presentation (i.e. a head mounted display, HMD) instead of a fixed monitor, see Kappé, Van Erp and Korteling (in press). However, the use of head-slaved camera control in combination with an HMD also has two potential drawbacks.
First, HMDs may influence comfort and control behaviour of the operator. Kotulak and Morse (1995) discuss a survey of 58 aviators by Behar, who found that 51% had visual discomfort, 35% had headache, and 21% had blurred vision. These symptoms could have a common origin: eye-head co-ordination could be affected by HMD characteristics, and smaller field sizes place heavy demands on head movements, since subjects must move their heads to sample the environment rather than using the more effortless joystick control. A study by Gauthier, Martin and Stark (1986) suggests that the greater head inertia associated with HMDs may induce a decrease in the amplitude-velocity relationship of head movements, i.e. slowing of head movement and small changes in head amplitude. Further, eye movements may change secondary to these changes in head velocity. Eye

movement maximum amplitude and velocity increase with increasing inertia. Gauthier et al. (1986) studied these effects of added head inertia and discuss that oscillopsia (continuous displacement or instability of the visual world) was prominent and consistent in perceptual reports of their subjects.
Second, transmission delays may distort the correct relation between the external environment and the perceived visual array. Because the images on an HMD are presented in the actual viewing direction of the operator, a transmission delay introduces a discrepancy between the viewing direction of the camera at the moment the images were recorded at the remote site, and the viewing direction of the operator at the moment the images are presented. This results in the operator perceiving the world as unstable when he moves his head. For example, when the operator has a steady image of an object, moving his head will 'drag' it across the environment during the transmission delay. Therefore, transmission delays will probably impede adequate spatial mapping of the visual information.
A possibility to reduce the first drawback (comfort) is to project the images in a moving window projected onto a dome, instead of on an HMD. A possibility to prevent the second drawback (delay) is to display the images in the viewing direction of the camera at the moment of recording, and not in the actual viewing direction of the operator (called *delay-handling* throughout the paper). This results in an image location which corresponds with the image content, and follows the actual viewing direction of the operator with a delay, instead of an image location which corresponds with the actual viewing direction, but not with the image content.

In case the field of view on the environment has the same size as the field of presentation (which is defined as the size of the display on which the view on the environment can be presented, e.g. the size of the dome), the principle of delay-handling will lead to image loss on the side contra-laterally to the direction of motion. Therefore, the field of presentation must preferably have spare space to overcome this loss. In this respect, domes are preferable. The size of this spare space and the transmission delay determine the maximum speed the camera can rotate without image loss.

Experiment

The present exploratory experiment was used to investigate the possibilities of head-slaved camera control for unmanned platforms. To elaborate on the possible drawbacks mentioned above, we used two *presentation modes*: a head-mounted display, and a moving window on a dome; and we introduced different transmission delays and tested the principle of delay-handling. To test the effect on the operator's sense of SA, we developed an experimental task, which included level one, two and three of SA as defined by Endsley (1995).

Subjects

Seven college-educated, right-handed male subjects (age: 20 to 27 years) participated in the experiments. All subjects had normal or corrected to normal vision, were paid for their participation, and had no experience with similar operator tasks.

Apparatus

All images were generated by a three-channel Evans and Sutherland ESIG 2000 image generator (30 Hz update rate). The images were presented via a head mounted display (N-Vision, 41.5° × 34.5°, 800×600 pixels H×V), or via a projection screen (a Seos PRODAS HiView S-600 projection system, consisting of a spherical dome and three video projectors; radius 2.9 m, 150° × 42°, 2400×600 pixels H×V). The subject's head was positioned in the centre of the dome. Head orientation (horizontal and vertical) was registered by a Polhemus Fastrack head-tracker (resolution 0.15°, 30 Hz), with the sensor coil either mounted on the HMD or on a lightweight plastic helmet (weight < 0.1 kg). Minimum delay between head-tracking and displaying was about 60 ms. Head tracker data was used as input for the mathematical model (ran with 30 Hz on a 486-based PC), which calculated the motions of the simulated (head-slaved) camera and the objects in the database. The mathematical model also simulated the transmission delay between the camera and the operator, by using a pipeline with a size of 30 times the transmission delay (s). A second 486-based PC was used for scenario generation and data storage (30 Hz sampling frequency).
The presented view on the environment (window) had a size of 13.3° × 10.0°, and could be projected in the actual viewing direction, or in the viewing direction of the camera for which the images were generated. Note that with a transmission delay this resulted in a delayed image content and a delayed image location, respectively.
The subject was seated in a chair with a right armrest, on which a spring-loaded joystick was mounted. A response button was mounted on top of the joystick (Figure 1).

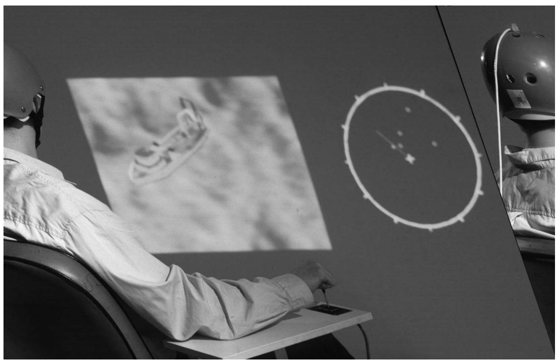


Figure 1: An overview of the TNO MUAV-simulator facility

Task

The camera-platform remained at a fixed position and orientation throughout the experiment, altitude of 500 feet. The virtual environment depicted by the camera

image consisted of a textured sea, twelve ships, and six square so called oil-rigs. The oil-rigs were arranged along imaginary gridlines, such that they enclosed an area defined by parallel and perpendicular lines between the rigs (Figure 2). This area was defined as forbidden for target ships. The distance between the platforms was 1000–2000 feet.

Six moving ships of equal type were defined as targets; the other six ships were distracters, were of a different, smaller type and had to be neglected. The targets moved at 45 feet/s along a winding route that was unknown to the subject, and had a maximum turn rate of 3°/s. The ships headed for an end position within the forbidden area.

Overall task instruction was to give a signal when a target ship entered the forbidden area, which actually consists of the following parts:

- determine the form and location of the forbidden area by detecting the position of the oil-rigs, and drawing imaginary borders,
- detect and monitor the position and track of the target ships,
- give a signal whenever a target ship enters the forbidden area.

This experimental task was designed to implement the different levels of SA as introduced by Endsley (1995). Level one refers to the position of the oil-rigs and the ships, their attributes, and their spatial relations in the environment. Level two refers to comprehending the significance of the different elements: which ships are targets, and which targets are heading for the forbidden area. Level three refers to the need to predict the future position of targets, e.g. assess which of the targets will reach the forbidden area first.

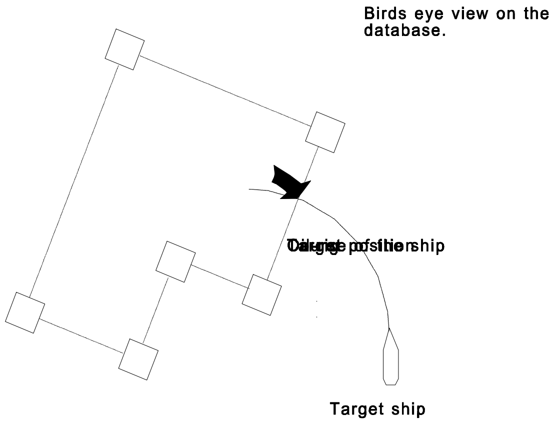


Figure 2: Illustration of a possible alignment of the six oil-rigs

At the time that one of the targets actually crossed a border (marked target position in Figure 2), subjects had to keep the ship’s stern in the centre of the camera image and push the button on the joystick. The target ship disappeared when it was held within 2° of the centre of the image at the time of the response. When the subject did not give a response, the target ship automatically disappeared when it reached a predefined end position within the forbidden area. Whenever a target ship disappeared, a new target ship was placed at a different position in the environment to keep the number of ships to be monitored constant during a run. A run was completed when six target ships had disappeared.

During the run, performance was recorded in order to calculate objective performance measures afterwards. Furthermore, after the completion of a session, subjects were given a post-test to ascertain that they had memorised the alignment of the oil-rigs, i.e. if they developed a mental model of the world during a run. A forced-choice procedure was used, in which the subjects had to choose the actual alignment of the oil-rigs out of the six drawings (bird’s eye view) of possible alignments.

Independent variables

Three independent variables were manipulated in a full factorial within subjects design: *presentation mode* (HMD and dome projection), *delay-handling* (absent, present), and *transmission delay* (0, 0.5, 1.0, 2.0, and 4.0 s), resulting in twenty conditions.

Dependent variables

The following performance measures were used:

- *Time to locate the oil rigs (s).* The measure was defined as the time it took a subject to locate all six oil-rigs, i.e. the time until the camera had been pointed at all of the six platforms at least once.
- *Time to border crossing (s).* The measure “time to border crossing” for each target was calculated as the time that a target was away from the border to be crossed at the moment of the response of the participant. Time to border crossing was taken over all targets signalled by the participant (between 1 and 6). This measure reflects the accuracy of the subjects in estimating the position, course and speed of the target ship relative to the oil-rigs, i.e. their accuracy in the perception and prediction of spatial relations.
- *SD heading (°).* The measure “SD heading” is defined as the standard deviation of the heading of the viewing direction during a single run, and is a measure of viewing behaviour.
- *SD pitch (°).* The measure “SD pitch” is defined as the standard deviation of the pitch of the viewing direction during a single run, and is a measure of viewing behaviour.
- *Multiple choice on platform orientation.* This measure was calculated as the number of correct choices of the alignment of the six oil-rigs (summed over the levels of transmission delay).

Statistical design

The experiment was completed in sessions consisting of the five transmission delay levels for a combination of presentation mode and delay-handling. These blocks of five runs were, although not completely, order-balanced across the subjects. Within each block, the order of transmission delay was randomised. For each subject,

the twenty scenarios were randomly assigned to the conditions, with the restriction that each combination of condition and scenario occurred only once throughout the experiment.

Each dependent variable was checked for outliers (scores that deviated by more than 3 SD from the overall mean) and sphericity. Incidentally, a large score on the time to border crossing was found. Target ships could approach a border until they were at a short distance from it, but because of the winding route they moved along, not actually cross the border. Therefore, values greater than 20 s were removed from the analysis. No other outliers were found.

Results of the performance measures "time to locate the oil-rigs", "time to border crossing", "SD heading", and "SD pitch" were analysed by a within-subjects design with three factors: presentation mode (2) × delay-handling (2) × transmission delay (5) with the statistical package STATISTICA 5.0. Significant results were further analysed by a post-hoc Tukey test. Results of the multiple choice question (only one observation per session of five runs) were analysed by a within-subjects design with two factors: presentation mode (2) × delay-handling (2).

Procedure

First, subjects received a brief written explanation about the general nature and procedures of the experiment. The instructor then showed the projection dome, chair, the plastic helmet and the HMD, and explained the purpose and task in more detail. The subjects came in pairs: one subject performed a session of five runs, preceded by a practice run, while the other subject rested. The practice run was with no transmission delay, was not registered, and performed with a scenario not used during the experiment. After a session the subject was instructed to perform the multiple-choice task in a room near the room in which the dome was situated.

Results

Presentation mode. On the basis of experimental observations (see Gauthier et al., 1986) and the smaller field of presentation, a disadvantage of the HMD was expected. However, none of the performance measures showed a significant effect of presentation mode.

Delay-handling. Two dependent variables showed a main positive effect of delay-handling. Time to border crossing showed a performance increase of 15% with the presence of delay-handling [means 5.8 s and 4.9 s, $F(1,6)=23{,}91$, $p<.01$]. The mean number of correct answers on the multiple choice task increases with 40% (means 2.4 and 3.4) with delay-handling present, $F(1,6)=21.00$, $p<.01$. Delay-handling showed no significant interactions.

Transmission delay. Three performance measures showed a main effect of transmission delay. The time needed to locate the oil-rigs [$F(4,24)=20.72$, $p<.01$], the time to border crossing [$F(4,24)=7.75$, $p<.01$], and SD pitch [$F(4,24)=6.39$, $p<.01$]. All effects showed performance decline with increasing transmission delay. The post hoc tests indicated that performance on the former two variables was degraded for delays larger than 0.5 s, on the latter only for a delay of 4 s.

Discussion

The present study concentrates on the concept of situation awareness (SA) in relation to camera control of unmanned platforms using virtual environment (VE) techniques. In the introduction, it was hypothesised that inherent characteristics of the man-machine interface, like the limited field of view and the time delay between image recording at the remote site and image presentation, may hamper the operator in developing a good sense of SA. Providing the operator with high quality information on (changes in) viewing direction by introducing a head-slaved camera system with head-slaved display may support the operator and improve SA. However, literature also shows that such systems may degrade other aspects, e.g. comfort, control strategy, and the spatial relation between viewing direction of camera and operator as a result of transmission delays.

The present experiment focussed on the applicability of head-slaved camera systems in MUAV applications. To overcome possible drawbacks of HMDs, we compared a head mounted display with a head slaved dome projection and to overcome the possible drawbacks of transmission delay. We introduced a mechanism of delay-handling which preserves the correct spatial relation between viewing direction of the camera and the operator by presenting incoming images in the camera viewing direction, and not in the actual viewing direction of the operator. A new experimental task was introduced to include the different levels of SA as discerned by Endsley (1995).

The results show no significant effect of presentation mode. Although mean values on SD heading and SD pitch showed higher values with dome projection over the HMD, the effects did not reach significance ($p=.16$ and $p=.10$, respectively).

The results indicated a positive main effect of the principle of delay-handling (depicting the delayed images in the camera, not in the actual head direction). Both the results of the time to border crossing and the multiple choice task show performance improvement when delay-handling is applied. Time to locate all oil-rigs and control behaviour did not differ with delay-handling absent or present. This indicates that delay-handling is especially useful for developing higher levels of SA, i.e. in determining the exact spatial relation between the oil-rigs and the imaginary borders and the targets.

The main effect of transmission delay shows that this variable both degrades the development of the sense of SA at all levels, and the control behaviour of the operator.

Because delay-handling results in a window moving with a delay, the available field of presentation must be larger than the field of view. This may be a disadvantage for the HMD mode of presentation, because HMDs have a restricted field of presentation. However, the lack of an interaction presentation mode × delay-handling shows that the field of presentation of the presently used HMD was sufficient.

We also expected an interaction between delay-handling and transmission delay. Increasing transmission delays will disturb the spatial relations more for the same control signals, and was therefore expected to increase the positive effects of delay-handling. Even a third order interaction (presentation mode × delay-handling × transmission delay) might have been present. Transmission delays were supposed to be compensated by presenting the images in the spatially correct viewing direction. This method requires a field of presentation, which is larger than the size of the camera images, and must be increased with increasing time delays. Since the field of presentation of the HMD is restricted, an additional advantage of the dome projection was expected for larger transmission delays. However, none of the interactions was found.

Recommendations

It is recommended to perform human factors research aimed at further improving operator performance by optimising interface design. Areas of interest include the following:

- Directly compare the effects of joystick versus head-coupled camera control on the sense of SA and camera control performance.
- Investigate the effects of a zoomed-in camera image on head-coupled camera control. The zoomed-in camera image disturbs the relation between head rotations and translational flow in the image, which may be confusing and uncomfortable to the operator.
- Further explore the applicability of the method of delay-handling in, for example, situations in which the camera translates through the remote environment, or in which the camera image is zoomed-in.
- Investigate the relation between man-machine interface characteristics and the different levels of SA, and develop specific operator support. An example is adding high quality visual information to the camera image to provide the visual information that is lost in some situations, e.g. as a consequence of the low update rate of the image (by presenting visual motion information), a zoomed-in image (by presenting correct translational flow for camera rotations), and transmission delays (by introducing a predictive display).

References

Barfield, W., Rosenberg, C. & Furness III, T.A. (1995). Situation awareness as a function of frame of reference, computer-graphics eyepoint elevation, and geometric field of view. *International Journal of Aviation Psychology, 5 (3),* 233-256.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37 (1),* 32-64.

Gauthier, G.M., Martin, B.J. & Stark, L.W. (1986). Adapted head- and eye-movement responses to added-head inertia. *Aviation, Space, and Environmental Medicine, 57,* 336-342.

Jeannerod, M. (1997). *The cognitive neuroscience of action.* Cambridge, UK: Blackwell Publishers Inc.

Kappé, B., Van Erp, J.B.F. & Korteling, J.E. (in press). Effects of head-slaved and peripheral displays on lane-keeping performance and spatial orientation. Accepted by *Human Factors.*

Kenyon, R.V. & Kneller, E.W. (1992). Human performance and field of view. *1992 SID International Symposium Digest of Technical Papers* (pp. 290-293), Playa del Rey, CA: Society for Information Display.

Kotulak, J.C. & Morse, S.E. (1995). Oculomotor responses with aviator helmet-mounted displays and their relation to in-flight symptoms. *Human Factors, 37 (4),* 699-710.

Schneider, G.E. (1969). Two visual systems. *Science, 163,* 895-902.

Smolensky, M.W. (1993). Toward the physiological measurement of situation awareness: the case for eye movement measurements. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 41-42).

Trevarthen, C.B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung, 31,* 299-337.

Thompson, J.A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance, 9,* 427-443.

Tyrell, R.A., Rudolph, K.K., Eggers, B.G. & Leibowitz, H.W. (1993). Evidence for persistence of visual guidance information. *Perception and Psychophysics, 54 (4),* 431-438.

Ungerleider, L. & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale & R.J.W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.

Van Erp, J.B.F., Kappé, B. & Korteling, J.E. (1996). *Visual support in target search from a moving unmanned aerial vehicle* (Report TM-96-A002). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Van Erp, J.B.F., Korteling, J.E. & Kappé, B. (1995). *Visual support in camera control from a moving unmanned aerial vehicle* (Report TNO-TM 1995 A-26). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Venturino, M. & Kunze, R.J. (1989). Spatial awareness with a helmet-mounted display. *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1388-1392).

This page has been deliberately left blank

Page intentionnellement blanche

The Dangerous Virtual Building, an Example of the Use of Virtual Reality for Training in Safety Procedures

Miguel Lozano[1], Marcos Fernandez, Joaquín Casillas, Javier Fernández & Cristina Romero
Institute of Robotics, University of Valencia
Polígono de la Coma s/n
MailBox: 2085 – 46071 — Valencia, Spain

Abstract

There is an ancient proverb that says *"Tell me and I will forget. Show me and I may remember. Involve me, and I will learn"*. This has been the main principle behind the big raising of immersive technologies in the field of training and education.

Here we explain our experience in using this kind of technology in the area of work risk and incident prevention. The high accident rate suffered by the construction sector has been one of the reasons that have moved us to develop the system that this article describes. The objective of the system is the training of the operator in safety procedures on the job. For this reason a VR system has been created that on the one hand reproduces a similar environment to the one experienced by the operator in real life, and on the other hand it provides for a number of operations to be completed. These operations which are very usual for the worker in real life imply a risk that must be understood by the worker, e.g. walking around the construction trenches carrying some type of load could cause a loosening of the ground resulting in death. For the complete training of the worker, the virtual environment contains the three fundamental phases of the construction of a building. Besides all of the general tools of the job may or may not have a safety component. So the number of dangerous operations that the system provides for and monitors are encountered in real life (working on a scaffolding, in trenches, on roofs, on the various floors, crashes, falls, overloads, etc.) By means of training and learning about the risks involved in the operations (from the most simple) you will obtain the best preparation in the sector, reducing therefore the rate mentioned above.

Using the system the worker is really involved in the task, and is able to understand the real risk that the task carries out, because he is in front of a screen that shows the object in its actual size and he has to make the proper decision. The system do not intent to train him or her in the skills of the task but in the safety way to proceed in its development.

This is a case that can be port to other military or civil areas where not only are important the skills but also is necessary to observe a methodology that ensures a safety performance.

We point out also in this paper how is possible using low-cost equipment to produce a good degree of immersive system. This is an important point in order to extend the use of those systems to such a sector or when the number of subjects to be involved in the training process make necessary to use a elevate number of simulation systems.

Introduction and capabilities of the system

Virtual environments are of major interest to computer graphics researchers; this is due, in part, to their ability to immerse the user in a computer-generated alternate reality in which we can easily recreate scenarios which are too dangerous, difficult or expensive to play in real life (Bukowski, 1997).

In this paper, we present an approach to this kind of system, the dangerous virtual building system (DVB) is an application of visual simulation, oriented to worker's education in the field of civil construction (Alkoc, 1993). One of the main goals to achieve by the system is the training of the operator in safety procedures on the job, and the second is to give us a measurement or an evaluation of these safety tasks.

The DVB has been designed to simulate a finite number of risky procedures that could occur in a real work environment. Demonstrating these procedures and evaluating the risks that each one implies, the workers can learn or review the safety routines that are often forbidden. Later the application will provide a measurement of the learning of capabilities of each worker in these safety procedures.

The main user in the DVB system, will be a student of a course in safety tasks, who works in construction field. Generally, the student doesn't know how to use most of the common tools utilised in computers such as a mouse or a keyboard. So in order not to hinder the learning process an instructor is needed to advise in the management of the system and to explain the goals to be achieved during the simulation.

A prototype of this system, based on SGI workstation, was developed and tested (Lozano, 1999) and currently we are developing a new open architecture focussed on the DVB training system.

The system under development consists of a centralised instructor control sever plus twelve simulation nodes (based on PC architecture). Each one of the subjects is immerse in his/her own simulation process and the

[1] For contact with the author: miguel@glup.irobot.uv.es

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

instructor can control the development of each exercise. The system has been based on distributed standard architecture (CORBA) and for the output of the simulation three possibilities has been offer: Head Mounted Display, flat monitor or 2×2 meters screen. The core of simulation graphics has been developed using low-cost graphics platforms with LINUX operative system and Performer libraries. The whole system offers enough graphic quality for both purposes, the training phase and the instructor node.

This instructor, who knows the capabilities of the system, will control the simulation, ordering some kinds of tasks for each student and enabling or disabling the proper conditions for that task. Later he will check the results given by the DVB in the learning process and will be able give us the level achieved for each student.

The way to show these procedures has been based on a VR application, so that, we can reproduce the familiar environment of the worker and he must interact with the system in order to achieve his goals. The input device used for the subject interact with the system has been and standard joystick.

The main capabilities of the system are:

- Simulate a virtual building environment managed from a subjective point of view (the camera) and controlled from the Joystick position (Cabral, 1996).
- Simulate and control the worker-environment interaction: The system simulates a number of risk situations (defined below), and must control the reactions and consequences.
- A number of elements (objects, tools, ... etc.) must be created and the interaction must be controlled by one specific module called the *worker's bag* (Santonja, 1996).
- The system takes into account the legislation regarding safety rules, and informs the worker if his behaviour doesn't comply with these rules.

In the rest of this article we will define each one of these capabilities, exploring in this way the contents of the system.

The Object Interface

The interaction with the objects commonly used in the building area is a very important element of the application. The application must allow the worker to be able to select objects and carry out an action with them. For this purpose, an object interface has been designed similar to the interfaces of adventure games. Whenever we wish we could show at the bottom of the screen an area composed of the next elements:

- The upper row is used to show the objects that the worker wears at a given moment, such as a helmet, gloves, etc.
- The middle row shows the objects that the worker carries in his hand, his pockets, or work belt, such a large hoe, a shovel, etc.
- The right area shows the objects that the worker carries in the wheelbarrow, if the worker finds it necessary to collect them. The wheelbarrow will then appear in the middle row as a transported object.
- The lower row shows the different actions that can be carried out with the currently selected object (Figure 1).

Figure 1: The Object Interface

Pressing the right spaceball button shows the object interface area. Once opened the object interface, can be in one of two possible states:

- Object selection: it allows the free choice between the objects that the worker wears, transport, or carries in the wheelbarrow.
- Action selection: it allows free choice between the possible actions for the currently selected object.

Once an object and the action the worker wants to perform on that object has been chosen, it will be checked to see if such an action is feasible with that object, and if it may carried out. The selection process can be summarised in Figure 2.

The verification of the action with the selected object is one of the most important steps in the diagram shown above. In order to carry out this verification, it is necessary is to take into account the weight and maximum volume the worker can support. Moreover, there is also a necessity to verify the number of 'spots' that remain free for an object to be carried. These 'spots' are the pockets, the belt, the worker hands, etc.

In order to help to the verification of the actions, a mask is assigned to each object with the possible action that can be performed with that object. As actions are performed over the objects, the mask will be modified to update the future possible actions over the object. In this manner the execution of an action over an object can be completely controlled.

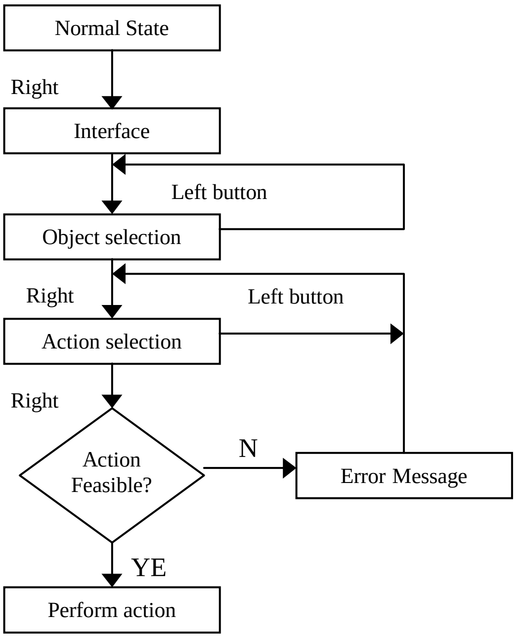


Figure 2: Object interface state diagram

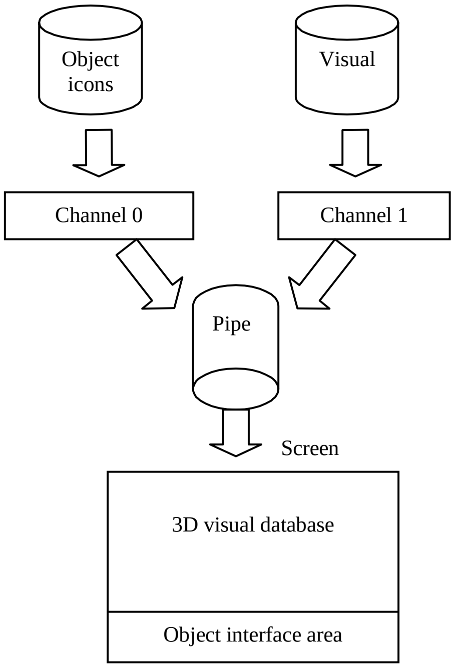


Figure 3: Channel Structure

In Figure 3 we can see an image of the application object interface with some of the building area objects loaded.

This figure shows us the objects that the worker wears in the first row of the object interface: the worker wears work suit, work boots, a tool belt, a safety harness, a helmet, and work gloves. In the row showing the transported objects, the worker carries an anti-gas mask and the wheelbarrow with the objects shown to the right of the interface (a large hoe, a shovel, a brick and a cement sack). The current selected object is marked with a blue square, as we can see in the helmet icon. The four actions possible over the selected object are shown in the lower row: cancel, transport, leave, and put into the wheelbarrow.

The interface object area is a dedicated channel different to the visual database scene channel, with its own visual database which is composed of small plane (two dimensional) objects with its texture applied (Rohlf, 1994). This structure is shown in Figure 3.

Danger situations

The main purpose of the application is training of the construction workers in issues concerning safety conditions (Lozano, 1998; Bukowski, 1997). This embraces knowing the essential equipment for each kind of task, and the right way of doing that task.

The stage has been divided in five areas and two access points, one for the workers and the other for the vehicles. The areas are arranged as shown in Figure 4.

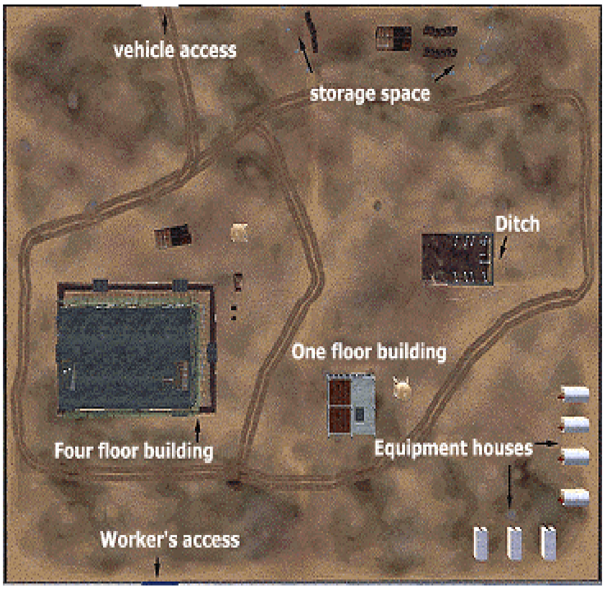


Figure 4: Arrangement of working areas

The five working areas are as follows:

- Equipment barracks: There are seven barracks with different equipment and clothes that the workers can use. There are elements, which are suitable for general working in the building area, such as boots and gloves. However, there is more than one element for each type of clothes. For example, there are rain and anti-slide boots, and the worker must choose the right equipment for the task he is going to do according to the weather conditions.

- Storage space: This is an area where the building elements are stored. The workers should leave things like cement sacks, wheelbarrows, and general working tools in this area.
- The ditch: There is a ditch with two propped up walls and the other two unpropped. The worker can go down to the ditch trough a ladder.
- One floor building: It is a small area with a building that has one floor and is under construction. There is a scaffold for the worker to use when working on the facade and a ladder to go up to the first floor.
- Four floor building: It is the biggest building in the building area, and is also under construction.

In addition, there are a couple of access points. The workers must use the people access point; otherwise they could suffer serious injury.

In the next images we can see situations corresponding to the areas mentioned above:

Figure 5: Working areas of the application

Training can be broadly defined as the learning or acquisition of skills in order to enhance performance at a given task or job (Burinston, 1995). In order to train, the building workers have thirty-six different dangerous situations that have been defined throughout the building areas. There are several types of situations:

- Situations that do not depend on the area which the worker is working in: this situations depends on time, altitude, weight, etc. Examples of these situations are:
  - Jumping: if the worker jumps from a surface (a scaffold), he could get injured if the altitude is moderate (up to four meters) or even die if he jumps from a higher site.
  - If the worker accumulate too many objects or materials in concrete areas, there is a risk of terrain collapse in those areas. So the worker must wear the necessary safety equipment in case such a situation occurs.
  - Collision with dangerous objects: if the worker drops a dangerous object (such a large hoe), it could cause injury to other workers. This will be advised by a warning message.
- Situations that depend on the place or area which the worker is working in: there are a lot of these situations, so we will describe a few organised by area:
  - Vehicle access area: if the worker goes into the building area through the vehicle entrance, he could be run over by a truck (Bayarri, 1996).
  - Storage space: here, the worker must walk carefully because there are dangerous objects in the area, so he should not stay on this area for a long time.
  - Ditch: in this area, there is a risk of terrain collapse if the worker walks in the zone that is not yet secure. If the worker wears a safety belt, he could be rescued in case of terrain collapse.
  - One floor building: here there is a scaffold that does not comply with to building regulations, so there is a risk of falling. The worker must be attached to the scaffold through the safety belt in order to prevent an accident.
  - Four floor building: in this area there are several different situations.

There is a trench surrounding the building with duckboards for the workers to go into. If the worker jumps the trench, he could fall and be seriously injured. So he must use the duckboards to access the building.

Walking under the building without a helmet is dangerous, because some object dropped from a higher floor may hit the worker.
There are gaps for the lifts that are surrounded by wooden fences, but in some cases the fence is not complete. In this case, the worker must pick up a wood board and complete the fence to prevent a possible accident (such as falling through the gap in the fence).
There are provisional ramps for the workers to go up to higher floors. Some of these ramps lack bricks, so the worker may slide and fall. In this case, the worker must use the correct ramps, or wear suitable boots.

The following images show some of the above situations:

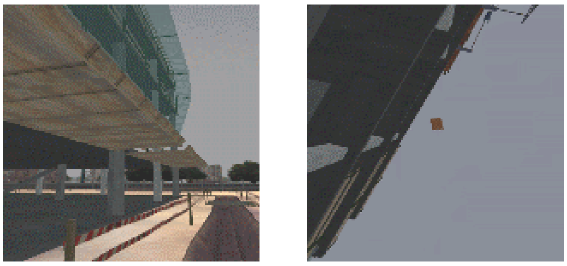


Figure 6: Risk of sudden fall of a brick from the roof

In the left-hand picture the worker is walking under a protective cover, so there is not a risk of the object falling. In the right hand picture, a brick is falling on to the worker. As there is no protective cover, the brick will hit the worker, and if he is not wearing a helmet he will be badly injured.

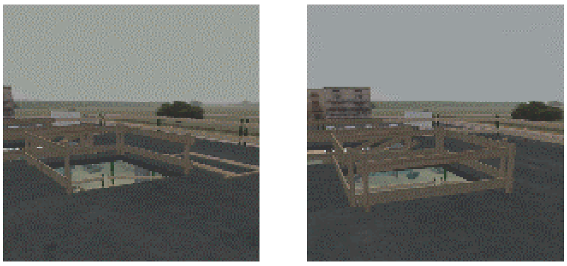


Figure 7: Risk of falling down through the lift gap

The left-hand picture shows that the fence surrounding the lift gap lacks an element. The worker must secure the gap zone by placing the wooden board on the floor.

In summary, the first action that the worker should undertake is to go to the equipment barracks and put on the correct clothes, depending on the area that he is going to work in. Then he will be prepared to begin his task, and go to the corresponding area. When finished, he must leave the elements, which he has worked on, in the storage area, and the clothes in the equipment barracks. In order to leave the building area; the worker must use the people access. The simulation restarts whenever the worker is killed by an accident, and the worker must go again to the barracks. In this manner the worker will learn the correct elements he must use in the corresponding areas.

Evaluation and future works

The previous prototype version of this system, which was running on a Silicon Graphics workstation, was tested with more than forty workers.

We can summarise the main objectives of those tests in two aspects: firstly, evaluate the degree of acceptance of the system in a group of people that is not familiar with this technology, and secondly evaluate transference of learning when it was produced.

Concerning to the first aspect pointed out previously, the basic analysis of the queries performed to the workers concluded that they were very excited with the use of this technology. At the first step users saw the system as a new experience and they were more active than in other teaching media, like video.
However at this point some problems were detected concerning to the navigation in the three dimensional environment and its location aspects. A couple of motion sickness cases were detected.
Focussing on the second point, a good learning transference was detected taking into account the written test performed after the exercises.

Nevertheless is important to notice that these tests were only initial evaluations that they only try to evaluate the convenience of starting the process of implementing actually a profitable system.

One of the problems detected in the first prototype was the high cost of the system. The current system has almost the same capabilities and a cost twenty times lower.

We have been also working in solving the problems of navigation, basically making the use of the joystick more intuitive and limiting in some ways the freedom of movements that some times produced a problem of location.

The current system as we have explained before is being developed with twelve simultaneous training nodes. The idea is to make this system portable in order to be installed into a forty-feed truck where the training process will be developed directly at the building area. By this way we will reduce the learning cost and will increase the productivity, taking into account the number of persons able to run the system.

The process of real evaluation of the system will start by the end of the year when the whole system will be ready to go to the real building areas.

References

Alkoc, C., “A computer Based Simulation of Concreting Operations”. Development in Civil & Construction Eng. Computing, 8. Publi. Civil Comp Press (UK). 1993.

Bayarri, S., Fernández, M. & Pérez, M., “Virtual Reality for Driving Simulation”. Communications of ACM (May 1996).

Bukowski, R. & Séquin, C., “Interactive Simulation of Fire in Building Environments”, SIGGRAPH 94.

Burinston, A. & Taylor, P., “Generic Training analysis methodology”, ITEC 95.

Cabral, B., “Graphics Techniques for walkthrough Applications”, SIGGRAPH 96, Course Notes.

Lozano, M., Casillas, J. & Fernández, J., “A Training System for Crane and Material Transport System Operation in Building Areas”, ITEC 98. Laussane. April 1998.

Lozano, M. Et al., “The dangerous virtual building, a VR training system for construction workers”, ITEC99. The Hague. April 1999

Santonja, J., Bayarri, S., Pérez, M. & Martinez, F., “Simplificación de mallas triangulares para su visualización en tiempo real”, CEIG’96.

Rohlf, J. & Helman, J., “IRIS Performer: A High Performance Multiprocessing Toolkit for Real-Time 3Dgraphics”, SIGGRAPH 94.

Visualisation of Geographic Data in Virtual Environments

Thomas Alexander[1]

Research Establishment for Applied Science (FGAN)

Research Institute for Communication, Information Processing and Ergonomics

Dept. Ergonomics and Information Systems

Neuenahrer Straße 20

53343 Wachtberg-Werthhoven

Germany

Summary

Virtual Environments (VE) are characterised as a computer-based generation of scenes of abstract or realistic environments, which can be perceived consistently. The use of VE is very promising in several areas, especially when visualisation of complex data in a realistic and clearly understandable way is needed. For military applications VE technology has potential in the area of research and development, training, mission support and mission rehearsal.

A further application is use in Command & Control (C²)-systems due to upcoming demands in this area.

In future battlespace scenarios huge amounts of highly dynamic information will be available due to the technical development of sensor, communication and information systems. Therefore advanced techniques for supporting the military commander and displaying complex tactical situation data in a clearly understandable way have to be developed and evaluated.

In this connection a concept for pre-processing and visualising incoming tactical data and three-dimensional geographical data has been developed. This “Electronic Sandtable (ELSA)”, as described in this paper, uses VE technology. The ELSA facilitates a plastic stereoscopic visualisation of three-dimensional data. It has been designed to simulate a sandtable as commonly used by the Armed Forces for tactical education and training.

Therefore the visualisation of digital geographic data (elevation (DTED) and feature (DFAD) data) is necessary.

This paper focuses on the stereoscopic visualisation of geographic data. Therefore different stereoscopic projection models are described and compared to each other. For the Electronic Sandtable a model with a window projection was chosen and implemented. The baseline concept and first results of this implementation are referred to in this paper.

1. Introduction

Huge amounts of highly dynamic information will be available in future battlespace scenarios because of the technical development of sensor, communication and information systems. Broad data acquisition, transfer, and presentation will enable the military commander to get a variety of diverse information about the battlefield scenario. The accomplished information dominance is more and more considered to be essential for a battlespace dominance.

But the massive quantity of information is also hazardous. Especially in time-critical situations when tactical decision making under stress is required, relevant information may be overseen and a wrong mental model of the tactical situation might be gained.

That overload is likely to be reduced by using new technologies for data pre-processing and data presentation. Because data presentation is of critical importance in the whole process of decision making, ergonomic research is required to analyse the whole process of data presentation, considering new displays and interaction devices.

Especially using Virtual Environment (VE)-technology is promising. It was found to have high potential in presenting and interacting with complex amounts of data. Therefore VE will increase the clearness and intelligibility of a complex tactical situation. The situation scenario is not perceived as a complex of abstract information but as a pseudo-realistic model landscape. This is intensified by an intuitive, easy to learn interaction with the included objects.

2. Command and Control (C²) Systems

Command and Control (C²) Systems have been designed to support the military staff in co-ordinating defensive, peace keeping and enforcing missions, exercises, humanitarian aid and ministerial expertise. For this reason diverse sensor information data and information data of knowledge databanks are joined in these systems.

A part of C² is the output and presentation of tactical information. It has large influence on the general decision making process, because the commander’s mental model of the battlespace situation is based upon the information perceived.

The SHOR-model (Stimulus, Hypothesis, Option, Response) of decision making introduced by Wohl (1981) proves this. It describes the process of decision making from data gathering to executing responses. The available and pre-processed information of a C²-System is displayed by the Tactical Situation Display (TSD).

2.1 Tactical Situation Displays (TSD) today

The basic function of TSD is to display the current situation of own and reconnoitred enemy troops and

[1] For contact with author: alex@fgan.de; tel. +49 228 9435 480, fax +49 228 9435 508

Paper presented at the RTO HFM Workshop on “What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?”, held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

facilities in the operation area to the commander of a military unit.
Moreover the TSD is used for tactical planning of intended future operations. Quantity and quality of situation data are essential for an adequate operation planning (Grandt et al., 1997).
Today's conventional TSDs might not be able to meet the demands of future battlespace scenarios and have to be extended by new, innovative technology. The strike forces today uses two basically different types of TSDs.
The first one, shown in Figure 1, is a *command post in the field*. The TSD used here works by means of paper & pencil. Actual information is transmitted by radio or field telephone and drawn into a map.

Figure 1: TSD at command posts "in the field"

It is obvious that in time-critical processes with large amounts of rapid changing information this leads to an overload of the operators. Moreover, the display may not show valid or actual information and causes errors in decision-making. However, it brings along the advantage that the commander is in the field: He gains high situational awareness, experiences the terrain, cover, weather, etc. and knows "what is really going on" at that place.
On the other hand there are *TSDs at operation centres*. Tactical situation data is pre-processed and computers are used to visualise the results.
The advantages of these computer-based TSDs are: Actuality of data, provided that the communication infrastructure is fast enough; and different views of levels of data aggregation and possibilities to include additional battlespace information.
But the flood of information may lead to an information overload and data representation is still limited to two dimensions and techniques of interaction with data have to be learnt.
The approach of using VE as TSD first expands the two-dimensional visualisation to three dimensions. This means that height information can easily be perceived. Additional elevation aids, like elevation profiles or colour texturing, can be skipped and replaced by others (e.g. reconnaissance photos, weather data, etc.
The more important thing is that general interaction with data is simplified and happens more intuitively. This facilitates an experience of the tactical situation and the generation of a correct mental model. In an ideal VE-system the computer is not realised as an active entity, but becomes an invisible assistant which knows about user intentions and supports him (Alexander et al., 1997). Therefore operator workload is supposed to be reduced and situational awareness to be increased.

2.2 Application of VE-Technology in C²-Systems

The amount of studies and applications in the area of VE and VE-technology has increased rapidly recently. But whereas VE is close to become applicable in research and development and for single training applications, studies considering the specific use of VE in C² have just begun. Therefore knowledge in this area is limited and a lot of projects are in a conceptual phase.
Most research studies and projects in this area have been started in the past two years. Because of ongoing development in this area this is only a brief overview. Detailed information is given in Alexander et al. (1999).
Generally speaking, the approaches can be divided into two groups. The first group consists of concepts and long-term programs including VE-components. This is a top-down approach, which is at high political level and typically application-oriented. The second group is characterised by specific VE-projects and laboratories. Consequently it follows a bottom-up approach and is presentation- and technology-oriented. Fortunately, there are links between both so that they meet and synergetic effects exist.
The Swedish *ROLF (Mobile Joint Command and Control System 2010)* is a long-term program. Its goal is to determine new possibilities for military commanders of using VE-Technology in mobile command posts. ROLF describes requirements for situational awareness, decision-making and support, work methodology and organisation of military crew and staff. The main idea is to use modern methods and technology to help a group of operators in difficult situations with complex, time-critical decision making. ROLF includes the *Aquarium* as TSD, which is a semi-immersive VE-system. The TSD is used to visualise positions of own and enemy troops, positions of important institutions, terrain and weather data in different views. Data pre-processing is used to select the data displayed and ensures that only important information is visible (Sundin, 1996).
Especially the realisation that in future battle scenarios all actions of the military commander will be in an unclear, vague environment and the importance of an information dominance led to the development of the *Command Post of the Future Program (CPoF)* of DARPA (1998). The program's goal is to accelerate the decision making process with ongoing reduction of the staff. Therefore new technology is needed to make maximum use of the whole human perceptory system in order to transmit maximum amount of information. This includes an interactive, three-dimensional visualisation, three-dimensional interaction with computer-generated objects, presentation of inaccuracy and probability, integration of dynamic factors, three-dimensional sym-

bolic, integration of natural language processing and integration of knowledge-based systems.

The second, more technology-oriented group of approaches is larger. Institutions and laboratories working in this area use different VE-technology. The technology is often reconfigured to be used in different research projects and experiments.

The *US Battle Command Battle Lab (BCBL)* performs conceptional studies as well as experimental analysis in a VR-laboratory. One goal is to develop a technology for multi-media, scene-based application in education and training for organisation and staff functions. This system shall be connected to the Internet to increase the range of application (Heredia, 1999).

At the *US Naval Research Laboratory (NRL)* an advanced battle planing and management system has been developed. The system works with a semi-immersive display and enables multi-modal interaction. It was found to be very suitable for virtual-prototyping, interactive mission planing and increasing situational awareness (NRL, 1997).

Similar approaches, like *Mirage* of the Army Research Lab (ARL) (IST, 1997), the *Visualisation Architecture Technology (VAT)* of the Crewstation Technology Laboratory (CTS) (Achille, 1998) or the *Electronic Sand Table* of MITRE Corp. (MITRE, 1998) also use a semi-immersive VE-technology, as described further on.

Other approaches use full immersive VE or desktop-VE respectively (Dockery & Hill, 1996; Morgenthaler et al, 1998).

2.3 The Idea of an Electronic Sandtable

The Electronic Sandtable at FGAN/FKIE has been developed as an advanced display for tactical information in mission planing, control and rehearsal. The concept is based on the sandtable metaphor. The military sandtable, as shown in Figure 2, consists of a sandy model landscape with simplified objects representing woods, buildings, points of interest or military units. It is broadly used in military education and training.

Figure 2: Sandtable in military education

But the traditional sandtable is static; all changes of deployment have to be done manually. Each change of region is very time-consuming and has also to be done manually. Moreover the accuracy for representing real geographic data is poor.

It is intended to model the sandtable by means of a VE-system. This way the system becomes capable of presenting dynamics, enabling real-time interaction and changes of the point-of-view while benefits of the real sandtable remain.

For this purpose geographic data and tactic data have to be visualised stereoscopically. It is intended to create a model landscape, in which dynamic battle scenario is included.

Furthermore additional functionality can be added, e.g. visibility, range of weapon systems, etc. The implementation of this idea will be described in detail in chapter 5.

3. Virtual Environments (VE)

The basic idea of generating and using a computer-generated artificial reality was mentioned first in science fiction literature at the middle of the 20th century. Due to rapid development of computer technology in the second half of the century, a partly realisation of this idea became possible. Nowadays these VE-Systems are commercially available and starting to be used for a broad range of applications (Alexander et al, 1999).

According to Bullinger et al (1997), Virtual Environments (VE) describe the computer-based generation of an intuitively perceivable and experiencable scene of a natural or an abstract environment. It is characterised by capacities for multi-modal, three-dimensional modelling and simulation of objects and situations. A further characteristic is the close interaction of the human operator with the system.

In this connection, Virtual Reality (VR) has been defined by NATO HFM-021 (nn.) as:

> *"... the experience of being in a synthetic environment and the perceiving and interacting through sensors and effectors, actively and passively, with it and the objects in it, as if they were real. Virtual Reality technology allows the user to perceive and experience sensory contact and interact dynamically with such contact in any or all modalities."*

This definition of VR, which is often used as a synonym to VE, overlaps with VE. But whereas VE is application oriented, VR describes, strictly speaking, a total model of the reality, including all manifold facets of it. As this is not possible today and may not be possible in future, the further article will be use the term VE.

VE-systems are on their way of becoming to be used for different applications. The main applications have found to be research and development, training, telemanipulation and teleoperations, mission support, and mission rehearsal. Further information about military applications is given in Alexander et al. (1999).

4. Geographic Data

Geography is the science of analysis of the surface of the earth and the earth-human ecosystem. The historic roots reach back to the antique world when geography was

used for the description of land, coasts and harbours. Still the description of the surface of the earth, called cartography, is one of the largest domains of geography. However, today geography is not limited to physics (geomorphology, climate, hydrographics, soil science, and geography of vegetation and animal), but includes political, social, economic and cultural aspects as well.
The structure of geographic databanks depends on the kind of application the data is intended for. Usually offices for land register and military offices are the main principals and users.
Data for military cartography has to be as exact, complete and actual as possible. This means a complete collection of data about all kinds of objects and the exact registration of their geographic co-ordinates is main criteria for structure of the referring databank.
The *geographic data* available is divided into (Helmuth, 1996):

- *Raster data*, which describes a subset of pixel data, like scanned paper maps of different scales. Assignment to other geographic data requires geo-referencing by means of the determined values for the map's corners.
- *Picture data*, which comprises geo-referenced or non-referenced aerial or satellite photos. Equalising reference points or procedures of aerial triangulation do geo-referencing.
- *Vector data*, which includes pre-processed data of surfaces (e.g. woods, lakes), lines (e.g. streets, rivers) or points (e.g. power poles, points of interests, bridges, towers) and the positions of their bases and attributes. Vector data is usually two-dimensional feature data and has to be merged with elevation data from other sources. For visualisation vector data is linked to detail objects.
- *Matrix data*, which describes terrain data structured and saved in matrix format. Usually, terrain data is organised like this.

All categories differ from each other in quality, resolution and actuality. Generally, data is available in scales between 1:25.000 and 1:250.000. The most common data-format is summarised in Figure 3.

Data Type	Name	Resolution / Scale (dep. on region)
Raster	MRG PCMAP ADRG	1:50.000 - 1:2.000.000 1:50.000 - 1:2.000.000 1:50.000 - 1:1.000.000
Picture	aerial photos satellite photos SPOT satellite photos Landsat-TM	1:32.000 & 1:70.000 10 m X 10 m 30 m X 30 m
Vector	DLM DCW DFAD VMAP U-VKN	1:25.00 - 1:1.000.000 1:1.000.000 & 1:2.000.000 1:250.000 1:50.000 - 1:250.000 1:50.000
Matrix	DHM/M745 DTED DGMA	30 m X 30 m 90 m X 90 m 90 m X 90 m

Figure 3: Different Formats of Geographic Data (Helmuth, 1996)

With growing demand on realistic education and training and ongoing technical development of displays new requirements for geographic data are emerged. In the future the main needs will be higher resolution and realistic texturing.
However, it cannot be taken as granted that all data required is available in the format, resolution and quality needed for the application. For this reason, an extension of one databank by different other databanks has to be done. This may lead to inaccuracies and inconsistency making further data processing necessary.

5. Electronic Sandtable (ELSA)

The Electronic Sandtable has been implemented as a testbed at the Research Institute for Communication, Information Processing and Ergonomics (FKIE). The structure and implementation of the semi-immersive VE-system is described in this chapter.

5.1 Baseline Structure

Because of the large size of geographic databanks and the need for real-time interaction, the underlying structure has been arranged in two stages (Alexander et al., 1997). A draft of this subdivision of the structure is given in Figure 4.

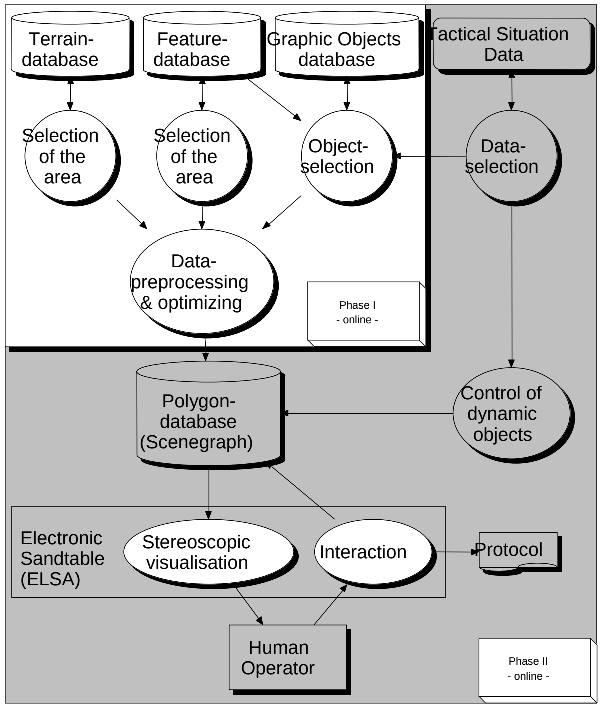


Figure 4: Structure of the Electronic Sandtable

The first stage is executed offline. In this stage the scene graph is determined. The scene graph is a hierarchically ordered databank of all polygons included in the visible scene.
In a semi-automatic process data and objects are selected, integrated and re-ordered with respect to maximum rendering performance. This re-ordered polygon-databank is called the scene graph. Afterwards

the scene graph stays constant without any changes of its structure.

In the second stage additional data is constantly added and the scene graph is visualised online. The additional data, i.e. tactical situation data and data from external data sources, is linked to objects of the scene graph. Additional input of external data using different protocols (DIS, HLA) shall also become possible in future. The incoming data controls position and status of military units. Additional data like actual situation videos or information of knowledge databanks can, also be included.

After that the rendering subsystem selects the visible subset of the scene graph. Out of this two separate projections are calculated and written into two-frame buffer. Then both frame buffers are visualised alternately on a horizontal plane.

The human operator interacts with the scene by means of different interaction devices. The inputs serve as commands, which affect the objects of the scene graph. They are logged for later analysis.

The operator is able to select different visible areas for navigation. The borders of the area serve as one input variable of the rendering subsystem. Additionally each of the operator's movements is tracked by a head-tracker. The position output of the tracker is another input variable of the rendering subsystem for new projection calculation.

5.2 Data Processing and Visualisation

For visualisation the geographic data has to be transferred into the scene graph to be visualised. The process is executed offline and done semi-automatically. It is divided into *data selection, pre-processing and optimising phase*.

In the first step an area of interest is selected and the relating terrain (DTED) and feature (DFAD) data is extracted. Additionally, links between features and geometric objects are defined. Afterwards the selected data is saved in a temporary buffer, which has to be pre-processed, and optimised for visualisation.

Geometric objects include the geometric description of the object (e.g. tanks, aeroplane) and additional information (e.g. unit status, damage reports, etc.). At the stage of real-time visualisation they are shown at the position given either by the geographic data or the tactical situation data.

The following steps of pre-processing and optimising are necessary because terrain and feature data are generated from geographic databanks. These databanks were designed with regard to different requirements, which makes them unsuitable for a real-time, realistic visualisation.

Pre-processing takes into account that consistency and integrity are highly important criteria for databanks. If datasets of more than one databank are merged, contradicting data might emerge and cause errors. Those errors are based on errors or inaccuracies in the original databanks, different data resolution or different actuality of data acquisition.

As soon as consistency and integrity is proved, the process of merging terrain and feature data starts. Geometric objects are appended and, if necessary, adjusted to ground level.

Finally the triangulation process starts and determines polygons for visualisation.

For real-time visualisation an optimising process has to be performed to keep the amount of rendered polygons minimal. Therefore the databank system transfers only information about the visual subset. Non-visible parts outside the field of view are clipped.

For further reduction the databank is re-organised and the scene graph is tiled. In the visualisation process the distance to the point of view sets the level of complexity for each tile.

Different levels of complexity called levels of detail (LOD) are another technique to reduce polygons. LOD means more than one representation of different levels of complexity (different amount of polygons) for the same subset. This means, if a subset gets closer to the point of view, a higher LOD with more polygons is visualised.

Using these techniques, data is re-organised with regard to visualisation issues. The output of this process is the scene graph, which can be visualised in real-time on the display.

5.3 Concept of Semi-immersive Display Technology

The display technology used for three-dimensional visualisation is a semi-immersive virtual workbench. Krüger & Fröhlich (1992) have originally developed this concept. The baseline concept is shown in Figure 5. Today it is used for various applications.

A projector projects two computer-generated, time-alternated pictures onto a mirror. The mirror reflects them to a horizontal focussing screen. By using shutter glasses, i.e. LCD-glasses shading each side alternately synchronous to the projection, the operators perceive two separate pictures for the right and the left eye. The synchronisation works by an emitter sending infrared signals synchronously to the picture projected.

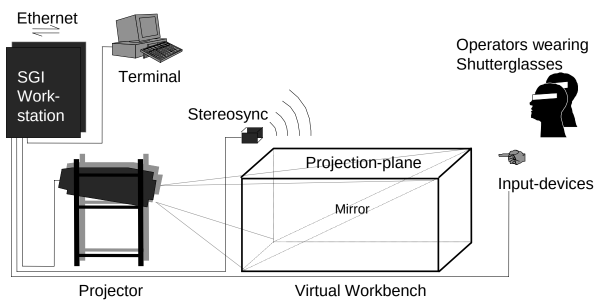


Figure 5: Principle of a Semi-Immersive Virtual Workbench

Finally, both pictures perceived are fusioned by the cerebrum to a single, three-dimensional model.

6. Stereoscopic Visualisation

The design of the user interface of VE-systems has been found to be one of the main criteria of quality for its application. The Electronic Sandtable (ELSA) serves as

the interface between the real environment on the one hand and the virtual scene on the other hand. Moreover it uses a different metaphor than the desktop-metaphor used in various computer applications. Therefore new interaction techniques and procedures have to be developed, analysed and optimised according to a high performance of the human-computer-system (Alexander, 1999).

A realistic, three-dimensional visualisation of terrain data has to consider the physiological procedures of visual depth perceiving. These procedures have been studied extensively, and several different hypothesis for depth perceiving exist.

Each hypothesis postulates the existence of depth cues. The classic depth cues will be summarised later in this chapter. Of those especially the stereoscopic disparity and parallax are of critical importance for the application of the Electronic Sandtable.

A computer-based visualisation has to take into account different depth cues. For stereoscopic visualisation different viewing models exist. The common models will be presented in this chapter as well.

6.1 Process of Visual Perception

The physiological visual system consists of the eye as sense organ for stimulus acquisition, the optic nerve for stimulus transfer and the optic centre of the cerebrum for stimulus processing.

According to Schmidt & Thews (1995) the human eye can be divided into two subsystems:

- Subsystem 1 performs the refraction of incoming light. Its main components are Iris (control of incoming light intensity), lens (refraction), vitreous body (stability) and diverse muscles (adjustment).
- Subsystem 2, jointly with the central nervous system, transfers the light to stimulus signals of nerve cells. It consists of the retina with its two different light receptors.

The stimuli are transferred via the optic nerve to the optic centres of the cerebrum. Here the optic sensing and recognition takes place.

Visual perception is generally based on three stages of perception (Kelle, 1994):

The first stage is an egocentric perception of the own person. This allows a separation of objects of the own body and other objects, making possible to determine the own position with regard to other objects and an *absolute depth perception*.

The next step is a comparison of the objects in the environment, allowing a *relative depth perception*.

Finally memory, experience and internal processing mechanism lead to *depth cues* being fundamental for spatial perception.

6.2 Depth Cues

Depth cues are visual system cues, which enable perceiving of spatial dependencies (Hodges, 1992; Schmidt & Thews, 1995). They can be divided into monocular and binocular cues.

Monocular cues are valid for perception with one eye only.

The main monocular cues are:

- *perspective*: The projection of three-dimensional environment onto a two-dimensional display surface has large influence on the subjective depth perception. Most common projection is the linear projection characterised by parallel lines meeting at a single vanishing point.
- *difference in size*: If same objects are shown at different sizes, the larger object seems to be closer than the smaller one. This criterion is basically a consequence of the perspective depth cue.
- *known dimensions of objects*: Know sizes of objects are also influencing the subjective depth perception.
- *shading*: Occluding and covering enables a perception of relative position of several objects with regards to each other. The object shown with a closed shapes is perceived as closer than the other.
- *light and shadow*: The shadow within an object makes conclusions about its spatial structure possible. Position and size of the outer shadow gives information about the kind of object (mountain or valley) and its size.
- *accommodation*: Examining and focussing an object requires an adjustment of the refraction of the optical lens to get a sharp picture on the retina. This is called accommodation.

The binocular depth cues require the total binocular eye system. They influence the perception of short to medium distances.

Traditional binocular depth cues are:

- *convergence*: For examining and fixation of a point with both eyes the eyeballs have to be counterrotated, until both lines of sights meet at the fixated point. Only if this happens the object is pictured at identical points of both retinas and a further processing of the stimulus is possible.
- *disparity and parallax*: If one object is focussed in space, other objects are represented at non-corresponding retina areas, causing two different pictures for the right and left eye. The disparity is defined as the distance between both single pictures. Because of the importance for the Electronic Sandtable, this depth cue will be described in detail in chapter 6.3.

Additionally to these static cues further dynamic cues exist which have large influences on the depth perception for medium distances (17–29 m) (Kelle, 1994). Because they are of no relevance for the semi-immersive display technology, they will not be described in this paper.

6.3 Disparity and Parallax

Disparity and parallax have a large influence on depth perception and are the main depth cues for stereoscopic visualisation. Therefore they are described more detailed.

The distance between both eyes leads to different representations of an object on the retina of the right and the left eye. Both eyes perceive the object with a

different perspective. The distance between both pictures is described by disparity.

If an object is looked at, it is represented at the fovea of both eyes. A round spatial surface exists (horopter), representing all objects on it on corresponding retina aerials. Objects at positions different from the horopter are represented at non-corresponding retina areas. If the distance from the horopter is not too large, the cerebrum fusions the right and left picture to a three dimensional model. If it is too large, disturbing double pictures are perceived (Schmidt & Thews, 1995).

Disparity is a mathematical dimension and cannot be determined practically. Therefore the dimension of the *stereoscopic parallax* has been introduced. For this a reference level has been used which is parallel to the eyes' level and runs through the fixation point.

Parallax has been defined as (Helmholtz, 1910, ref. in: Kelle, 1994):

$$p = b_a \times a \times \frac{t}{e * t + e^2}$$

p = parallax
b_a = inter ocular distance
a = distance eyes / reference level
e = distance reference level / object
t = distance eyes / object (=a+e)

Parallax is also a dimension for depth separation and depth perception. Therefore it is deduced that depth perception decreases with square distance. Furthermore it increases linearly with inter ocular distance.

According to Kelle (1994), stereoscopic disparity and parallax has been found to be useful only for near and medium distance (maximum of 6–9 m).

Visualisation of geographic data of large scale means a large distance between eye point and surface. It can be concluded that exact modelling means that parallax and stereoscopic depth perception will be very low. Consequently, an exclusive use of real values for the model parameters (e.g. depth scale) would lead to no stereoscopic depth perception and the scene would be perceived flatly. On the other hand, too large values e.g. for depth scale would make the terrain more mountainous and may cause a wrong mental model of the terrain. For an ideal depth perception these parameters have to be adapted so that operators perceive the terrain structure subjectively correctly. Therefore a dynamic adaptation of the interocular distance of operators and depth scaling is needed.

Pilot experiments for determining optimum inter- ocular distance and depth scale have just started.

6.4 Stereoscopic Projection

For three-dimensional stereoscopic visualisation three different projection models are commonly used. Their baseline geometry is illustrated in Figure 6.

In Computer Aided Design (CAD), aerial photo analysis and for head-mounted-displays (HMD) *projection models with parallel line of sights* are used, as shown in Figure 6 (a). They are based on the assumption of a centre eye-point perpendicular to the projection plane. Right and left projections are calculated by using offset values and parallel shifting the projection right and left. The disadvantage of this model is that the scene can only be visualised underneath the projection plane. This is inconvenient for the concept of the Electronic Sandtable, because the scene would always be located beyond hand range. Another disadvantage is clipping at the borders of the display as missing visual information for either right or left eye appears. Especially at large displays this is very irritating for operators.

Figure 6 (b) shows the geometry of a *projection model using rotated line-of-sights*. Here the projections are rotated in the way that both lines-of-sight meet in the projection plane. The lines-of-sight do not stay perpendicular to the projection plane. It enables a visualisation underneath and as well as above the projection plane. There are no irritating effects on the borders of the display either. But because of the special geometry, an error of vertical parallax occurs. It can be observed at the borders of the display, where both lines meet at a point, which is above the projection plane. This leads to a "winding"-effect and the scene seems to be projected on a cylinder rather than a plane. Vertical parallax has found to be irritating especially on large displays.

The last projection model uses *window projection*, which means that two windows are introduced through which the virtual scene is perceived. The windows are positioned in the same level as the projection plane. Both lines-of-sight meet at the projection plane and remain perpendicular to it. In this model, stereoscopic parallax is only dependent on the distance to the display and no vertical parallax is introduced.

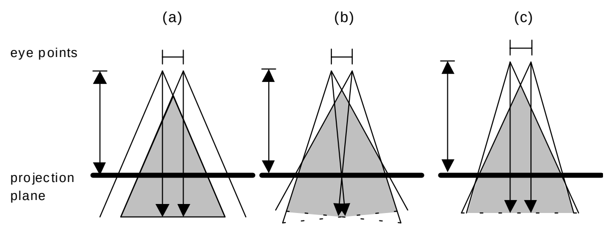


Figure 6: 3 projection models: (a) parallel projection, (b) rotated projection, (c) window projection

This model is used for the Electronic Sandtable. As shown in Figure 7, an asymmetric pyramid describes the model for each eye. This means, the perpendicular line through the top does not meet the centre of the pyramid basis.

For each projection six parameters are used to identify the pyramid. They include the values for front, back, top, bottom, left and right clipping plane. These values are calculated by x,y,z-position of both eye-points, scale factor and the display size as input.

Pilot experiments have shown good results for this projection model. Only little perspective error due to tracking of real eye position was determined. In future, this error will be minimised by calibrating the tracking equipment.

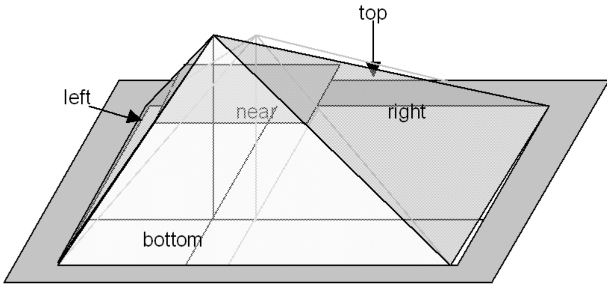


Figure 7: Right and left asymmetric projection pyramid and boundary surfaces (clipping planes)

7. Conclusion and Future Research

In this paper the baseline concept of using semi-immersive VE-technology as advanced TSD has been described. The approach has been shown to be promising and advantageous.

It has been emphasised that human factors and ergonomics are the main issues for reasonable VE-application. In this paper some research issues were introduced and results of ongoing research studies in the area of visualisation were presented.

So far only real-size shapes have been visualised. In future geographic data of different scales will be used. To evoke a stereoscopic depth perception, an adaptation of the scale factor for elevation as well as the dimension of inter ocular distance is necessary.

Another research topic is the maximum vertical range of the display. The display technique causes contradicting depth information, because both eyes accommodate on the projection plane, but fixate an object closer or more far away. However, if the virtual scene is too close, parallax becomes too large and the cerebrum cannot fusion both pictures. Therefore another research topic will be to determine the maximum useful vertical display range and the variability of human sense perceiving.

Pilot experiments in the visualisation area have been started and are currently going on.

Other important areas with high influence on the applicability of VE in C² are interaction and co-operation.

Interaction with the databank means navigation in the scene and manipulation of virtual objects. Procedures (software) and interaction devices (hardware) have to be designed, evaluated and analysed according to the application for both subgroups.

The concept of the Electronic Sandtable has been designed to enable multiple operators working in the virtual scene. It has to include co-operation concepts. In contrast to full-immersive VE, in semi-immersive VE all operators are present at the same location. Communication and inter-operator interaction work the natural way. Therefore mainly human-computer interaction issues have to be analysed. These main issues and problems are the development of a general concept for co-operation and co-operation procedures.

But even if in future the system works as it is supposed to be, one question to be answered still remains: The question for quantification of the profit and gain of using VE-systems. The key criteria for answering this question will be performance of the human-VE system.

For this reason human performance metrics will have to be introduced, formulated and analysed. They should be as fundamental as possible, but still take into account the characteristics of the application.

Jointly with other basic research studies they will be the key issues of future research in this area.

References

Achille, L.B. (1998): *Naval Air Warefare Center Aircraft Division Crewstation Technology Laboratory*. CSERIAC, Vol. IX: 2 (1998). Wright-Patterson AFB: FRL/HEC/CSERIAC, USA.

Alexander, T. (1999): Virtual Reality als Koordinations-werkzeug bei Planungsaufgaben. In: Gärtner (ed.): *Ergonomische Gestaltungswerkzeuge in der Fahrzeug- und Prozeßführung. DGLR-Bericht 99-02.* Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt e.V., pp 51-63.

Alexander, T.; Grandt, M.; Gärtner, K.-P. (1999): *Virtuelle Umgebungen für die Gestaltung zukünftiger Lagedarstellungssysteme*. FKIE-Bericht Nr. 4. Wachtberg-Werthhoven: FGAN-FKIE.

Alexander, T.; Wappenschmidt, A.; Gärtner, K.-P. (1997): Ein Verfahren zur Darstellung von Geodaten in Echtzeit. In: Möller (ed.): *Tagungsband 5. Workshop Sichtsysteme – Visualisierung in der Simulationstechnik*. Aachen: Shaker Verlag.

Bullinger, H.-J.; Brauer, W.; Braun, M. (1997): Virtual Environments. In: Salvendy (ed.): *Human Factors and Ergonomics*, pp 1725-1759. New York: John Wiley & Sons.

DARPA (1998): *Project „Command Post of the Future“*. http://www.darpa.mil/iso/cpof/main.htm

Dockery, J.T.; Hill, J.M. (1996): Virtual Command Center. *Proceedings of the 1996 Command and Control Research and Technology Symposium*, Sn. 61-68, Washington, D.C.: ACTIS.

Dockery, J.T.; Hill, J.M. (1996): *Virtual Command Center*. Proceedings of the ICCRTS 96. Washington D.C.

Grandt, M.; Alexander, T.; Gärtner, K.-P. (1997): Virtual Environments for the Simulation of a Tactical Situation Display. In: Holzhausen (ed.): *Advances in Multimedia and Simulation. Human-Machine-Interface Implication. Proceedings of the Europe Chapter of the Human Factors and Ergonomics Society Annual Conference.* Bochum: University of Applied Sciences.

Helmholtz (1910): *„Handbuch der physiologischen Optik“*. Verlag von Leopold Voss, Hamburg.

Helmuth (1996): Überblick über die verfügbaren MilGeo-Daten. *Wehrtechnisches Symposium „Die digitale Karte“*. Mannheim: Bundesakademie für Wehrverwaltung und Wehrtechnik.

Heredia, M.D. (1999): *Battle Command Battle Laboratory Fort Leavenworth, KS: Overview*. http://cacfs.army.mil/.

Hodges, L.F. (1992) : *Time multiplexed stereoscopic computer graphics*. IEEE Computer Graphics & Applications, pp. 20.

Institute for Simulation and Training (IST) (1997): „*Mirage - A New Way o View the Virtual World*". http://www.ist.ucf.edu/labsproj/projects/mirage.htm.

Kelle, O. (1994): *Dynamische Tiefenwahrnehmungs-kriterien in computergenerierten interaktiven Szenen und virtuellen Simulationsumgebungen*. Düsseldorf: VDI-Verlag.

Krüger, W.; Fröhlich, B. (1992): *The responsive workbench (virtual work environment)*. IEEE Computer Graphics and Applications, Vol. 14(3), pp 12-15.

MITRE (1998): *The Electronic Sand Table*. http://www.mitre.org/pubs/showcase/virtual_sand.html.

Morgenthaler, M.; Steiner, G.; Mayk, I. (1998): The virtual command post. *Proceedings of the 1998 Command and Control Research and Technology Symposium*. Monterey: Naval Postgraduate School.

NATO HFM-021 (n.n.): *Proceedings of the Workshop on Virtual Reality and Military Applications. 5.-9.12.1997, Orlando, USA.* Neuilly-sur-Seine: NATO RTA.

Naval Research Laboratory (NRL) (1997): „*Virtual Reality Responsive Workbench for Sea Dragon*". http://overlord.nrl.navy.mil/workbenchinfo.html.

Schmidt & Thews (1995): *Physiologie des Menschen*. Springer-Verlag, Berlin, Heidelberg, New-York, pp 278-315 (Gesichtssinn und Okulomotorik).

Sedgwick, H.A. (1986): Space and Motion perception. In: *Handbook of Perception and Human Performance, Vol. 1*. New York: John Wiley & Sons.

Sundin, C. (1996): ROLF — Mobile Joint C2 System for the year 2010. A vision under development and test. *Proceedings of the 1996 Command and Control Research and Technology Symposium*, Sn. 176-187, Washington, D.C.: ACTIS.

Wohl, J.G. (1981): *Force Management Decision Requirements for Air Force Tactical Command and Control*. IEEE Transactions on systems, man, and cybernetics, Vol. SMC-11 (9), pp 618-639.

This page has been deliberately left blank

Page intentionnellement blanche

Acquiring Distance Knowledge in Virtual Environments

Prof. Dr. Edgar Heineken[1] & Frank P. Schulte[2]
Gerhard Mercator University
Lotharstraße 65
47048 Duisburg
Germany

Abstract
Experimental results on the perception and cognition of distances in virtual environments are reported. These results show differences in the accuracy of distance perception depending on whether they are presented in desktop- or HMD-VR. In addition, they show that distance cognition in virtual environments is based on online-judgements (perception based) or on inferential judgements (memory based) depending on the subject's goal when navigating through the environment. Without an explicit goal to learn distances (incidental learning condition) the estimated length of routes in a virtual environment is inferred by the number of features (feature-accumulation-hypothesis) experienced on the respective route, just like in natural environments.

1. Introduction

A spatial environment can be explored directly or by means of a map. A number of studies dealing with the acquisition and representation of and the access to spatial information have documented differences in spatial learning associated with different modes of experience. Direct and map experience lead to a different understanding of the environment. Navigating through an environment enables subjects to estimate route distances and route orientations (route knowledge), whereas Euclidean distances and locations of landmarks (survey knowledge) are easier to estimate if the environment is presented using a map (Thorndyke & Hayes-Roth, 1982; Giraudo & Pailhou, 1994; Taylor & Tversky, 1996).

Spatial cognition research is becoming increasingly interested in the use of virtual environments as experimental settings: virtual reality technology provides both an economical and flexible design of realistic environments as well as a reliable registration of the subjects interactions with the environment.

The results of spatial cognition research are of practical interest when virtual environments are used as visualisation or training tools. Thus the question arises, whether there are differences in processing spatial knowledge (landmark-, route- or survey-knowledge) in natural and virtual environments (Wilson, 1997; Witmer, Bailey, Knerr & Parsons, 1996; Ruddle, Payne & Jones, 1997; Rossano, West, Robertson, Wayne & Chase, 1999).

The paper refers to the acquisition of distance-knowledge in virtual environments, and to the perception and cognition of distances.

In chapter 2 an experiment designed to compare desk-top and HMD-VR with respect to supporting distance perception is presented.

In chapter 3 a series of experiments on distance *cognition* in virtual environments are reported.

Chapter 4 summarises and discusses the results presented.

2. Distance Perception and Perceived Depth in Virtual Reality

There are different kinds of psychological spaces. A vista space means a space up to 30 m, explored by looking ahead without locomotion. This kind of psychological space can be contrasted with the environmental space (the entire space is not visible from the starting position, it can be explored only by locomotion), and the geographical space (the space is so large, that it can be explored only by means of a map). When designing vista spaces in virtual reality factors determining human space-perception have to be considered. There are nine different sources of information the human visual system uses as depth cues: occlusion, height in the visual field, relative size, relative density, aerial perspective, binocular disparities, accommodation, convergence and motion perspective (see Cutting, 1997). It is of interest, however, whether the perceiver's kind of interaction with the virtual environment (e.g. whether the view's orientation changes depending on the user's head movements, or not) may also affect their spatial sensitivity and thus their perception of distances in the space.

The hypothesis that distance perception in a virtual vista-space is more accurate in HMD-VR than in desktop VR is tested in a bisection-experiment.

A total of 18 subjects (7 male and 11 female) participated in the experiment. Their average age was 26 years, ranging from 20 to 36 years. The environment used in this experiment was created and presented using Superscape VRT 5.50 software, running on a 500 MHz Pentium III PC equipped with 196 MB RAM and a 32 MB Matrox G400 graphic accelerator card.

The environment showed a small forest through which a path led. The whole scene consisted of 8000 facets, and the maximum frame rate was limited to 20 frames per second to avoid lag differences. The environment was rendered in 640 to 480 pixel resolution.

Half of the group of subjects experienced the virtual environment by means of a head-tracked HMD

[1] For contact with Prof. Heineken: Heineken@uni-duisburg.de; tel. +49 203 379 2541, fax +49 203 379 1846
[2] For contact with Schulte: Frank.Schulte@uni-duisburg.de; tel. +49 203 379 2519, fax as above

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

(Virtuality Visette Pro combined with a Polhemus InsideTrak), the other half viewed the environment as a video projection (JVC DLA 10 SXGA). It was made sure that the FOV in both VR conditions was identical. In both conditions the subjects remained in a standing position.
The subjects were instructed to bisect a route presented to them in the virtual environment by moving a marker to the mid of the route. Figure 1 shows the virtual environment.

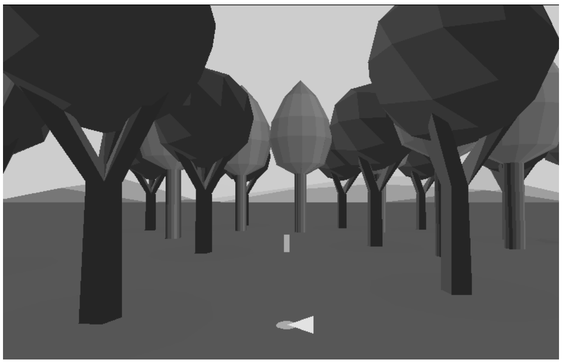


Figure 1: Starting point (circle) and end point (bar) of the presented route. The marker (triangle) has to be moved to the mid of the route. (note: ground texture has been deleted for printing reasons)

The presented routes differed with regard to their length — short (approximately 150 cm) or long (approximately 600 cm) — and with respect to the starting position of the marker (above the mid and below the mid). Each route hat to be bisected four times by each subject, twice in ascending and twice in descending order with respect to the initial position of the marker. The subjects stood 400 cm away from the route's starting point.
The participants in each experimental group were given different amounts of time to explore the virtual environment before their bisection task (30 seconds, 60 seconds and nor opportunity).

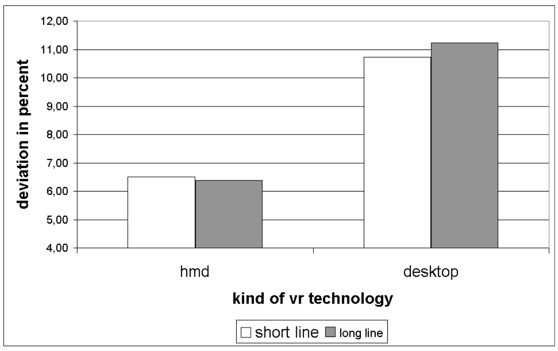


Figure 2: Bisection error (deviation from the real mid in percent) for the short and the long route under different VR-conditions (HMD, desktop)

The error of bisection was calculated as the absolute mean differences between the estimated mid and the real mid. Figure 2 shows that the bisection error is greater when the line is presented in desktop VR than when it is presented in HMD-VR. The difference is statistically significant ($F(1,12)=4,92$, $p<.05$). The factors "route length" and "experience with the virtual environment" did not affect the bisection error.
The results are showing that immersion improves depth perception and facilitates the judgement of distances in a virtual vista space.

3. Distance Cognition in Virtual Environments

There are two conflicting theories which try to predict the cognition of distances experienced in environmental spaces: the Feature-Accumulation-Theory (Sadalla & Magel, 1980) and the Route-Segmentation-Theory (Allen, 1988). According to the first theory, the cognitive distance of a route is positively correlated with the number of features experienced on the route, whereas the second theory proposes a positive correlation between estimated distance and the number of segmentations of the route.
Within the scope of those theories on distance cognition a series of experiments have been realised by Petra Jansen-Osmann in our institute. In the following section we report the main results of her doctoral thesis (Jansen-Osmann, 1999).

3.1 Number of route turns and estimated route length

The length of a route with more turns is estimated longer than a route with fewer turns. This result of an experiment carried out by Sadalla & Magel (1980) was replicated in a virtual maze. 20 subjects navigated successively through two mazes. The routes were of same length but differed in the number of turns (2 turns, 7 turns). Afterwards, they had to travel on a straight route until the distance covered seemed equal to the route travelled in the maze.
The covered distance was significantly de-pendent on the number of turns ($t(1,14)=3,56$, $p<.005$). The route with more turns was estimated longer (Figure 3).

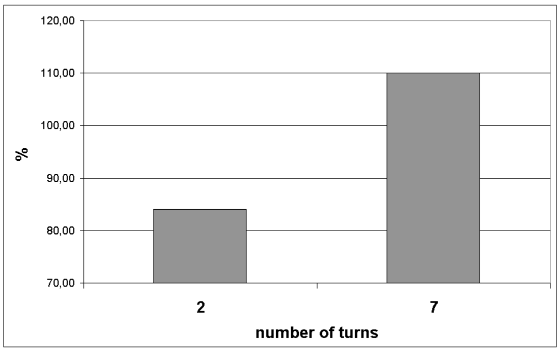


Figure 3: Cognitive distance (magnitude estimation) for routes with different numbers of turns

The result corresponds with both hypotheses on distance cognition because turns can on the one hand be regarded as features and on the other hand as borders of route segments, i.e. as segmentations.

3.2 Feature accumulation or route segmentation as determinants of distance cognition

In an experiment with 30 subjects distance estimations for segmented routes, routes enriched with features and empty routes were compared.

A street-scene was used in desktop-VR. On each side of the street 9 identical looking houses were presented. The number of houses and the number of crossways could not seen by the subjects from the starting point (Figure 4).

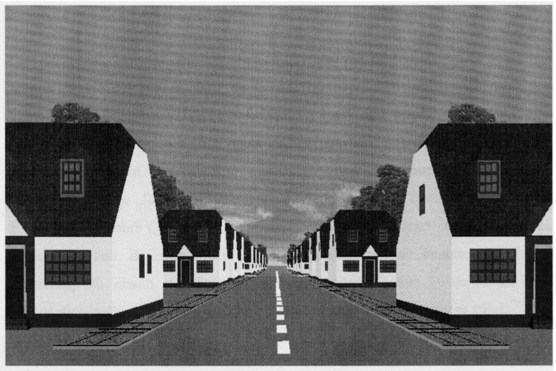


Figure 4: User's view at the starting point (note: crossways cannot be perceived at the starting point)

The spacing between the houses (Figure 5) as well as the location of crossways was varied. The subjects had to estimate six different distances between houses.

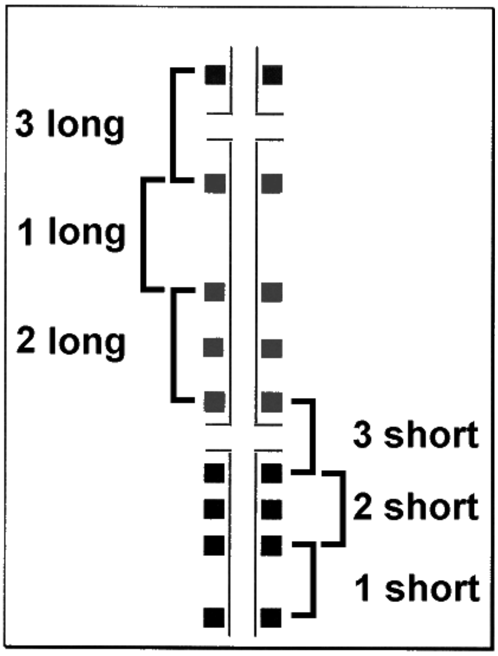


Figure 5: Empty sections (1), filled sections (2) and segmented sections (3) of different length in a survey map of the street scene used in the experiment

The whole route could be broken down in empty sections (1), sections filled with a house (2) or sections segmented through a crossways (3). Each kind of section could be short or long. Half of the group of subjects navigated through the virtual street using a joystick (active navigation), the other half experienced the street without joystick (passive navigation). Both groups experienced the environment three times successively.

Afterwards the subjects had to collocate the 9 houses on a vertical line on a sheet of paper with respect to their respective distances (Figure 6) and is consequently overestimated.

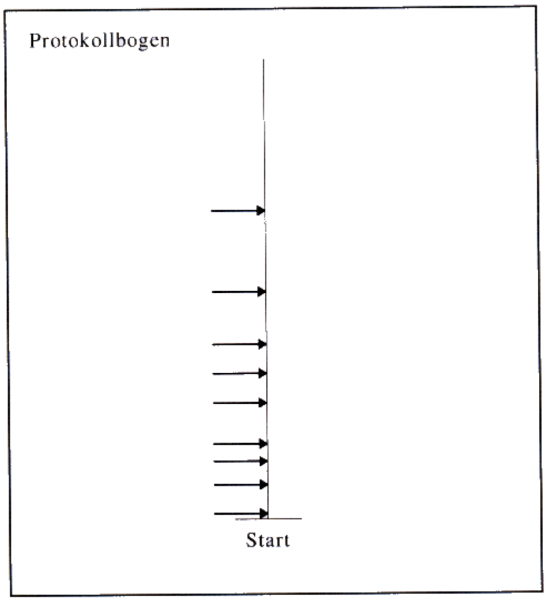


Figure 6: Protocol sheet: Collocation of the houses with respect to their distances

The results show that both the segmented and the filled sections were estimated equally longer than the empty sections of the route, and that this difference was more pronounced when the subjects had actively explored the virtual environment (Figure 7).

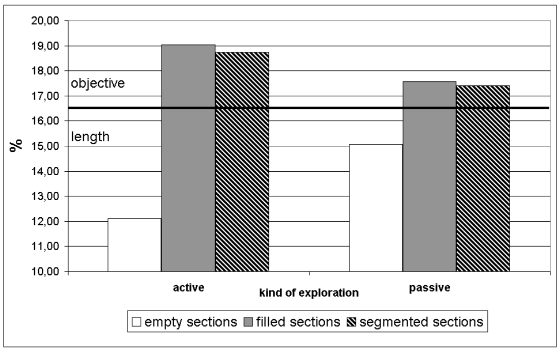


Figure 7: Cognitive distances (estimated length in relation to the total route length) of empty, filled and segmented route sections experienced actively or passively

Only the effect of the route design (empty, filled, segmented) on the distance estimations ($F(2,56)=12,16$, $p<.001$) was significant, showing that feature accumulation as well as route segmentation determine distance cognition in virtual environments.

3.3 Distance cognition based on the presentation of a survey map

Survey maps of the virtual streets used in the last experiment were presented to 15 subjects on a monitor for 1 minute after they had actively explored the virtual environment. Their distance judgements were clearly dependent only on route segmentation and not on feature accumulation (Figure 8).

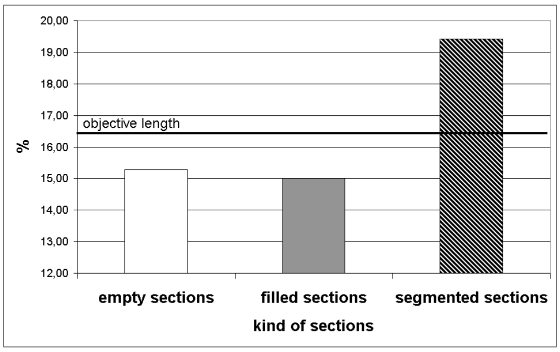


Figure 8: Cognitive distances (estimated length in relation to the total route length in percent) of empty, filled and segmented route sections experienced in a map

The route design significantly influences the distance estimation if the environment is presented on a map (F(2,28)=8,73, p<.01) but only the route segmentation — and not the feature accumulation — determines the perceptual organisation of the map and as a consequence the distance cognition (see Figure 9). When the street scene is presented as simultaneous structure the distance between two houses segregated by the crossways is perceptually strengthened.

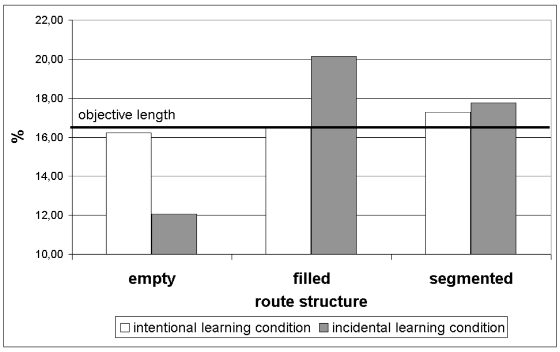


Figure 9: Cognitive distances (estimated length in relation to the total route length in percent) of empty, filled and segmented route sections in the case of incidental and intentional learning

3.4 Online- vs. inference-based distance judgement

The learner's goal when navigating through a virtual environment is a crucial encoding-factor in the processing of distances. In an experiment 30 subjects navigated through the same virtual environment used in the last two experiments. Half of them were instructed that afterwards they would have to estimate distances, whereas the other half was not explicitly instructed to focus on the distances.

There is a systematic interaction between the factors "route-design" and "kind of learning" (F(2,56)=11,36, p<.01) indicating that route-segmentation or feature-accumulation determine distance cognition only in case of incidental learning. If distances are learned intentionally, which means that the subjects encode distance directly, features and segmentations have no effect on the distance estimation: the distance estimation is based on the perceived distances (online judgement). In the case of incidental learning distances are not encoded directly, they are inferred afterwards using houses or crossways as heuristics (inference-based judgements).

4. Concluding Remarks

Accurate distance perception and distance cognition are necessary for applying VE in the field of training and are therefor a prerequisite for its validity as a training tool.
There are differences in the accuracy of distance perception depending on whether the environments are presented in desktop- or HMD-VR: immersion improves depth perception and facilitates the judgement of distances in a virtual vista space. Obviously the perceiver's sensumotoric interaction with the virtual environment provided by the tracking system enhances his spatial sensitivity.

Distance cognition in a virtual environmental space can be based on online-judgements (perception based) or on inferential judgements (memory based) depending on the subject's goal when navigating through the environment. The learner's goal is a crucial encoding-factor in the processing of distance-information. It determines the kind of spatial knowledge transferable from the virtual to the natural environment.

When VE are applied in the training of real world skills based on accurate distance perception and cognition, the designer should be familiar with psychological factors which determine the learner's spatial encoding and judgement of distances (e.g. the role of feature-accumulation). It was shown that without an explicit goal to learn distances the learner stores general information (features) when navigating through the environment, and later on judges the distance of a route by using the frequency of features experienced on the route as a heuristic.

References

Cutting, J.E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments & Computers, 29 (1)*, 27-36.

Gillner, S. & Mallot, H.P. (1998). Navigation and Acquisition of Spatial Knowledge in a Virtual Maze. *Journal of Cognitive Neuroscience, 19,* 445-463.

Giraudo, M. & Pailhous, J. (1994). Distortions and fluctuations in topographic memory. *Memory and Cognition, 22,* 14-26.

Jansen-Osmann, P. (1999). *Kognition von Distanzen — laborexperimentelle Untersuchungen in virtuellen Umgebungen.* Dissertation. (http://www.uni-duisburg. de/diss/diss9906/)

Ruddle, R.A., Payne, S.J. & Jones, D.M. (1997). Navigation buildings in desktop virtual environments: experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3*, 143-159.

Sadalla, E.K. & Magel, S.G. (1980). The perception of traversed distance. *Environment and Behavior, 12,* 65-79.

Taylor, H.A. & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371.391.

Thorndyke, P.W. & Hayes-Roth (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14,* 560-589.

Tlauka, M. & Wilson, P.N. (1996). Orientation-free representation form navigation through a computer-simulated environment. *Environment and Behavior, 28,* 647-664.

Wilson, P.N. (1997). Use of virtual reality computing in spatial learning research. In N. Foreman & R. Gillet (Eds.), *A Handbook of Spatial Research Paradigms and Methodologies.* Vol. 1. East Sussex, U.K.: Psychology Press, pp 181-206.

Witmer, B.G., Bailey, J.H., Knerr, B.W. & Parsons, K.C. (1996). Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human Computer Studies, 45,* 413-428.

This page has been deliberately left blank

Page intentionnellement blanche

Development of Virtual Auditory Interfaces

LCDR Russell D. Shilling[1], MSC, USN
United States Air Force Academy
Colorado Springs, CO 80840
USA

Tomasz Letowski, Ph.D.
Army Research Laboratory
Aberdeen Proving Ground, MD 21005
USA

1. Introduction

The design of visual components in virtual environments has shown rapid improvement and innovation. However, the design of auditory interfaces has lagged behind. Whereas visual scenes have become more compelling, the auditory portions of VE remain rudimentary. This disparity is perplexing since auditory cues play a crucial role in our day-to-day lives. Imagine entering a meeting with a room full of people. When you enter the room, you realize that the speaker's voice is emanating from all points in the room, yet the room is totally anechoic. In addition, you see other attendees moving in the room, yet there are no additional noises in the room except the speaker's voice. Despite walking into a "real" environment, your sense of reality would most probably be challenged. In fact, it is generally believed that the sense of presence is dependent upon auditory, visual, and tactile fidelity (Sheriden, 1996). Although the sense of realism in VE is also dependent on visual fidelity, virtual or spatial sound has been shown to increase the sense of "presence" (Hendrix, 1996). It stands to reason that when we develop poor auditory interfaces in a VE, the perceived quality of the entire VE is compromised (Storms, 1998). The problem with audio is that our normal auditory environment is "transparent". We don't consciously process a sound in our environment unless we NEED to attend to it. Yet, when slogging through mud while on patrol, soldiers use auditory cues to keep track of the people around them while scanning for threats in front of them. They don't need to keep looking at the people around them. While not consciously processing the sounds of their comrades, if someone stops walking, they'll recognize the lack of sound instantly.

2. Methods of Sound Presentation

There are a variety of ways to present sound in virtual environments. The most traditional method is to use speakers to present sound either monaurally, in stereo, or in surround sound. Speaker systems are bulky, do not typically provide elevation cues, and do not allow the sound engineer to have complete control of the auditory environment. Speaker systems DO allow for the possibility of presenting auditory stimuli such that the entire body is stimulated, especially when powerful subwoofers are employed. On the other hand, using headphones in conjunction with signal processing techniques, it is possible to generate stereo signals that contain most of the normal spatial cues available in the real world. Spatialized audio uses actual pinna cues stored as Head Related Transfer Functions (HRTFs) to give the perception of auditory objects as completely externalized in azimuth and elevation (Wightman & Kistler, 1989; Begault & Wenzel, 1993). When coupled with a headtracking device, spatialized audio provides a true virtual auditory interface. Using a spatialized auditory display, a variety of sound sources can be presented simultaneously at different directions and distances. One of the early criticisms of spatialized audio was that it was expensive to implement, however, as hardware and software solutions have proliferated, it has become feasible to include spatialized audio in most systems. Spatialized audio solutions can be fit into any budget, depending on the desired resolution and number of sound sources required. Most head-mounted displays are currently outfitted with headphones of sufficient quality to reproduce spatialized audio, making it relatively easy to incorporate spatialized audio in an immersive VR system. A complete lexicon for understanding and developing auditory displays can be found in Letowski, Vause, Shilling, Ballas, Brungert & McKinley (2000).

3. Effects of Auditory Displays on Performance

Illustrating the importance of sound, research conducted using spatialized auditory displays has demonstrated the importance of spatialized auditory cueing for reducing response time in cockpit applications. Spatialized auditory threat and attack displays were designed and implemented for both the pilot and co-pilot gunner in an AH-64 simulator at the Army Research Institute at Fort Rucker, Alabama (Shilling & Vause, 1999; Shilling, Letowski, & Storms, 2000). In this application, a ground-to-air missile display was supplemented with a spatialized auditory cue corresponding to the actual location of the missile relative to the pilot and co-pilot gunner. Figure 1 shows the difference between spatialized and normal displays for the response time to make the first 5 degrees of turn away from an incoming threat. Response time was reduced by approximately 350 msec. These data are consistent with previous research which demonstrated that response time to visual targets was significantly reduced when paired with a spatialized auditory stimulus (Perott et al., 1991) and the latency of saccadic eye movements was reduced when using spatialized auditory cues (Frens, Opstal & Willigen, 1995). In this same manner, auditory cueing can be used to compensate for the effects of limited FOV HMDs (Shilling, 1996). Applications can be further supplemented by exaggerating normal auditory cues

[1] For contact with author: Shilling85@alumni.wfu.edu

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

through so-called "supernormal localization" (Durlach, Shinn-Cunningham et al., 1993). Finally, using spatialized sound, speech intelligibility can be improved when applied to multi-user virtual environments and multi-channel radio communications (Haas, Gainer, Wightman, Couch & Shilling, 1997).

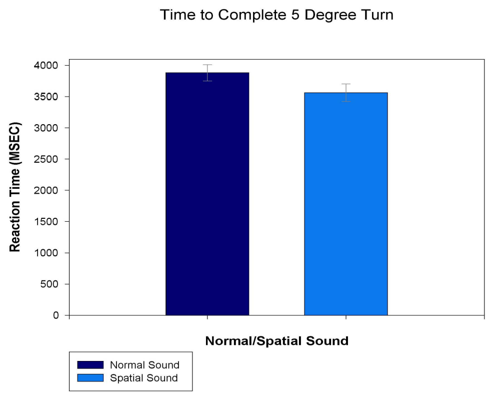


Figure 1: Difference between spatialized and normal displays

4. Lessons from the Entertainment Industry

The entertainment industry has recognized the importance of sound processing for over a century and has learned many important lessons that can be applied to problems in VE. At the beginning of the century, the Edison Standard Phonograph represented the cutting edge in audio technology. The method for cutting grooves in the wax cylinders was robust and resistant to the effects of scratches. However, consumers soon abandonedwax cylinders with vertically etched grooves for the less robust wax platter with horizontally etched grooves, because the platters were easier to store. Today, even though we have the technology to create astounding audio when developing VE's, it is more convenient to ignore the auditory interface because customer's aren't "requiring" high quality audio, software applications are not typically easy to implement, and the contributions of high quality sound are more subtle than for visual cues.

For instance, in motion pictures, sound has long been recognized as playing a crucial role in the emotional context of a film. Current efforts in my research are focusing on applying lessons learned from the film industry to problems associated with sound quality and emotional content in VE. Much can be learned about auditory special effects and sound system design from Hollywood. The first real attempt at immersing the audience in sound occurred with the production of Disney's "Fantasia" in 1939. Disney's sound engineers created a system called "Fantasound" which wrapped the musical compositions and sound effects of the movie around the audience. Though not a stereo production, the effects were quite astounding. However, the system required massive amounts of vacuum tube electronics and 54 speakers spread around the theater at a cost of $84,000 per theater. Virtually no theaters invested in the system and "Fantasound" was never used again. Today, we have a similar problem with applying sound in VE. Although the cost of consumer audio equipment has rapidly increased in quality and decreased in cost, systems designed for VE's are currently expensive and the development software to implement them is limited. Spatial audio sound servers, for example the AuSIM Acoustetron and the Tucker-Davis Technologies PD-1, typically cost in excess of $12,000. High cost and limited software availability are clearly the result of a lack of competition in audio products for VE.

5. Systematic Approach to Sound Design

On the practical side, the problem is not with the software engineers as much as with the lack of a clear set of requirements for implementing sound in VE. What is needed is a systematic approach to rendering the auditory environment necessary for any given application. When we want to render visual scenes, we rely on film as a reference. Unfortunately, when we design auditory scenes, we typically rely only on memory. In my laboratory, I am currently attempting to develop a systematic approach to cataloging the auditory environment to give the software engineer an objective reference to compare the sound in the VE with the real world experience.

One of the current efforts in my lab is to develop a systematic approach for obtaining baseline data concerning the content of an auditory environment. In addition to cataloguing the different sounds in a real environment, it is also important to systematically measure the intensity of sounds being experienced by the listener. In this manner, the VE developer has a highly detailed reference with which to compare the real world auditory environment with the virtual auditory environment. Two systems are currently being evaluated. The first system uses a portable Sony TCD-D8 DAT recorder coupled with Sennheisser microphone capsules (Figure 2). The microphone capsules will be inserted into an observer's auditory meatus (ear canal). In this manner, a complete spatialized recording can be made of the auditory environment, completely externalized with azimuth and elevation cues. The second system (Figure 3) is more robust, using a larger set of microphones produced by Core Sound which can clip to a set of eyeglasses to produce a binaural recording, complete with interaural time and intensity cues. Although, pinna cues cannot be utilized, the advantage of the latter system is that it would be more tolerant of extreme conditions, especially if the recordings are made outdoors. Both systems can be clipped to the belt and will be used in conjunction with a real time logging and event analyzer (CEL 593). The complete data set including sound recordings and sound measurements will be stored on CDROM for ease of use. The digital recordings also allow for spectral analyses to be conducted on specific auditory stimuli contained on the tape so that synthesized versions of those stimuli can be constructed.

Figure 2: The used portable Sony TCD-D8 DAT recorder coupled with Sennheisser microphone capsules

Figure 3: The set of microphones produced by Core Sound

References

Begault, D.R. & Wenzel, E.M. (1993). Headphone Localization of Speech. *Human Factors, 35 (2),* 361-376.

Durlach, N.I., Shinn-Cunningham, B.G. et al. (1993). Supernormal auditory localization. I. General background. *Presence, 2 (2),* 89-103.

Frens, M.A., Van Opstal, A.J. & Van der Willigen, R.F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Pereption & Psychophysics, 57 (6),* 802-816.

Haas, E. C., Gainer, C., Wightman, D., Couch, M. & Shilling, R.D. (1997). Enhancing System Safety with 3-D Audio Displays. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting.* 868-872, Albuquerque, NM.

Hendrix, C. & Barfield, W. (1996). The sense of presence in auditory virtual environments. *Presence, 5 (3),* 290-301.

Letowski, T., Vause, N., Shilling, R., Ballas, J., Brungert, D. & McKinley, R. (2000). *Human Factors Military Lexicon: Auditory Displays* (ARL Technical Report, ARL-TR-in print). Aberdeen Proving Ground, MD: Army Research Laboratory.

Perott, D.R., Sadralodabai, T., Saberi, K. & Strybel, T.Z. (1991). Aurally aided visual search in the central visual field: effects of visual load and visual enhancement of the target. *Human Factors, 33,* 389-400.

Sheriden, T.B. (1996). Further Musings on the Psychophysics of Presence. *Presence 5(2),* 241-246.

Shilling, R.D. & Vause, N. (1999). Tri-service cooperation, a key aspect of ARL-HRED spatial audio research. *Focus Army Research Laboratory, 5 (5).*

Shilling, R.D., Letowski, T. & Storms R. (2000). Spatial Auditory Displays for use within Attack Rotary Wing Aircraft. *Proceedings of the International Conference on Auditory Displays,* April 2000, Atlanta, GA.

Storms, R.L. (1998). *Auditory-visual cross-modal perception phenomena.* Doctoral Dissertation. Naval Postgraduate School, Monterey, California.

Wightman, F.L. & Kistler, D.J. (1989). Headphone simulation of free-field listening. II. Psychophysical validation. *Journal of the Acoustical Society of America, 85,* 868-878.

This page has been deliberately left blank

Page intentionnellement blanche

Educational Conditions for Successful Training with Virtual Reality Technologies

Dr. Alexander von Baeyer & Dr. Hartmut Sommer
IABG[1]
Einsteinstraße 20
85521 Ottobrunn
Germany

Summary

The paper focuses on those pedagogical conditions, which should be met, in order to assure successful training using virtual reality (VR) technologies. Therefore, neither new technical inventions nor large scale technical experiments are the issue of this paper. Instead a systematic catalogue of pedagogical questions will be proposed, which should be answered, before virtual reality is planned for training purposes.

The pedagogical catalogue is derived from the basics of educational psychology and media didactics. It comprises

- a taxonomy of learning objects, which are most suitable for virtual reality
- an analysis of training strategies and methods, as to how well they are suited for training in an almost entirely synthetic environment
- an analysis of the transfer of training, when virtual reality is the major training medium
- and finally rules and basic cost data, which may help to conduct cost effectiveness analyses.

Introduction

In this paper I will try to give a short and comprehensive overview on the basics of educational theory, which should be applied to training with VR technologies. I will do this in five statements. Each statement or thesis is accompanied by explanations. I start with a new look on a well known definition.

Probably everyone in this conference knows, what VR is. Nevertheless, I will give my own add-on to a commonly used definition and comment this definition. I do this, because I want to define important educational issues.

The common definition reads as follows:

VR is “a multi-dimensional human experience which is totally or partially computer generated and can be accepted by those experiencing the environment as consistent” (NATO DRG Panel 8 on Human Sciences, RSG 16).

My add-on is:

VR is a capability beyond life, virtual and constructive simulation and of course much beyond Computer Based Training systems, however it can be coupled with CBT. VR can be created, in order to convey training objectives and support training strategies.

Basic Statements

1. Statement

If training is the aim of VR, VR training programmes must comply with the basics of social and educational psychology.

These basics do not differ from what should generally be valid about training with constructive and virtual simulation. VR is an other example that there should be such things as simulation didactics. VR, however, increases the pedagogical requirements to be considered. These requirements concern mainly

- the distribution of learning material in a multi-sensory (multi-channel) experience (e.g. seeing, hearing, feeling of one’s own body, feeling of material properties, stress, decision making)
- the real experienced presence of an instructor and of other students during the learning and exercising process (social learning)
- the merging into VR and leaving the virtual environment (e.g. different feeling of own security).

Related to these three general problems are the following practical questions, which will partially be answered in this paper:

- Are VR technologies justified by relevant *training objectives*?
- Do VR training programmes enhance the quality of instruction and bring about better *training strategies*?
- Can the typical military *crew and leadership behaviour* be preserved in VR, where this is necessary for training?
- Are the offerings of VR *accepted* by experts of training and operation as an environment that facilitates learning?
- Will there be a chance to construct a *consistent* training scenario with new synthetic elements of the human environment?

These are the educational questions, which the VR community is invited to discuss further.

2. Statement

If we take the classical taxonomy of learning objectives, VR can be a relevant medium in complex psycho-motor training, only for certain cognitive tasks, may be to indoctrinate in the emotional and affective domain and (as a still controversial matter) in a real social context.

In principal VR is useful for the following four types of non-trivial application:

[1] For contact with authors: baeyer@iabg.de and hsommer@iabg.de, respectively; tel. +49 (0)89 6088 2127, fax +49 (0)89 6088 3612

Paper presented at the RTO HFM Workshop on “What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?”, held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

- Perceptual-motor learning, where real images are mixed with virtual components, e.g. the real hand manipulating computer generated interfaces (this is also called Augmented Reality),
- Perceptual cognitive training, when it becomes necessary to build a "mental map" on the basis of experience from various sense channels, not only based on the visual system, e.g. complex assembly tasks involving orientation in space, finding objects and moving them from one place to an other, discriminating different objects
- In general for team training in large scale exercises like C² training, large staff exercises, disaster control, but only as far as co-ordination skills and procedures are concerned
- And finally the exploration of unknown environments, provided that the data are up to date.

Examples for these types of application are

- Mission rehearsal, where all merits of VR are combined
- Reconnaissance, where VR however must have an added value to conventional simulation and training.

The total immersion into a synthetic environment leads to the exclusion of non-intended and disturbing information. This fact can be used or better: misused for indoctrination purposes. Sales promotions, radical behaviour changes, rapid conveying of emotional stimulus response patterns can be the objectives of such techniques. This again leads to the question, if and how much VR inhibits the ability of critical distance to the learning of those tasks, which require a critical attitude, e.g. all tasks comprising decision making between not fully transparent alternatives.
The impact of total immersive VR technology on the emotional behaviour is therefore a challenging new research question.

Social learning is however not yet sufficiently researched in fully immersive VR. The main problem lies in the isolating effect of VR. This means that it is still a not yet proven hypothesis, whether the acquisition of interpersonal skills, even and especially if they are interconnected with cognitive or procedural tasks can be supported by those VR technologies, which isolate the individual from direct personal contact with another individual in the same learning group. There are, however, semi-immersive VR-technologies like the cave- technique or the virtual workbenches, where individuals interact with each other "naturally". These techniques cover therefore in principle the all classes of learning objectives.

3. Statement
Training Strategies in VR do not differ much from those in virtual simulation and in CBT. However, they require more dedicated analysis and development, because VR offers more perceptual cues.
In comparison to constructive and virtual simulation VR has some distinctive features, which make it particularly valuable for articulated teaching and learning strategies. These features are:

- a broader perceptual spectrum
- a higher degree of differentiation in the perceptions (e.g. more depth cues)
- a higher degree of interactivity with the virtual environment.

These three properties of a deeper immersion into the artificial world offer the possibility, to differentiate and structure learning activities in a more effective way.
The advantages of learning and teaching with VR technologies are:

- *more learning material can be presented to the students*
- part task and part function training can be applied to a broader variety of learning tasks
- feedback control of learning success can become more differentiated and apply to a broader spectrum of tasks
- it may become easier to compose a set of part tasks to a real world like whole task in a almost realistically perceived learning environment.

However, VR requires a much more developed art of constructing the curricula and of designing the learning programmes and the learning aids. In short: VR makes the training development much more demanding and requires higher developmental qualifications.

4. Statement
The transfer of training into the operational situation has to be carefully analysed, because VR represents nevertheless only a part of "real reality".
As we have already said, the social dimension of reality is still hardly present in learning with VR technologies. Along with this, decisive other aspects of the learner are still drastically altered. These are

- the perception of the bodily self, which may be necessary in many psycho-motor learning tasks
- the unnatural feeling of wearing a helmet or a glove, which does either not resemble the normally worn helmets and gloves, or is a totally unrealistic feeling
- the multi-sensory perception of the environment. e.g. the not real feeling to walk a distance
- the apperception of the partner in the learning process, whenever this may be required for the acquisition of team building skills
- the apperception of the instructors, whenever this may have a motivational effect on the learning process or is a part of team building skills — remember that in typical military tasks training and personal example and leadership can not be separated.

All this means that skill acquisition by means of VR technologies puts the learner in sometimes extremely artificial surroundings, encapsulates his consciousness and lets him leave this virtual world with a repository of artificial behaviours. The first thing after leaving the artificial world of VR is to re-learn those behaviours, which do not fully comply with the operational

environment, to de-condition the learner away from the partially reduced and partially enriched experience towards a normal interaction with the operational environment. This again means, that although VR is an expensive training and an often time valuable medium, the transfer of training cannot be taken for granted and must be ascertained with much effort. If the curricular and didactic analysis has identified those tasks and skills that cannot be trained with VR, the transfer of training of the remaining VR-prone tasks can be evaluated without too big problems.

5. Statement

Cost and effectiveness of training with VR must be compared with training using virtual simulation. Whenever virtual simulation is feasible, VR should be analysed, whether it can produce better or cheaper solutions than virtual simulation.

On the effectiveness side of the comparison cost effectiveness analyses should consider the following issues:

- The enhanced representation of new and extended sensorial perceptions may *increase* the effectiveness.
- The possibility of mission rehearsal and procedural training in extreme situations, where total immersion is the only realistic experience, may also *increase* the effectiveness (good example may be the training for operations and maintenance in space or deep water).
- The reduced personal and interpersonal experience is definitely a factor, which *decreases* the effectiveness of VR in training.

On the cost side of the comparison the following issues should be considered:

- The HMD technology is a cost *decreasing* factor.
- The software development is a drastically increasing factor.
- Re-training and special transfer of training analyses can become cost *increasing* factors.

Therefore, considering VR for training should always start with cost effectiveness analyses based upon thoroughly conducted training analyses. However, the cost savings can reach several orders of magnitude, if training using VR is correctly designed. Examples are cargo handling skills or air drop skills, where the real aeroplane would be too expensive and virtual simulation is not giving the necessary depth cues.

Conclusion

To conclude this survey: What are the conditions of success of VR in training?

1. For the time being a limitation to tasks, which do not imply any personal proximity of other persons.
2. For the future more critical research into the interpersonal and social impact of VR and how far social interactions can be simulated in an total immersive environment.
3. Always limitation to empirically researched and proven simulation cues.
4. Always embedded in a well controlled transfer of training evaluation.
5. Always planned on the basis of cost effectiveness analyses.

This page has been deliberately left blank

Page intentionnellement blanche

Entertainment Technology and Military Virtual Environments

Michael R. Macedonia[1]
US Army Simulation, Training, and
Instrumentation Command
12350 Research Parkway,
Orlando, FL 32826, USA

Paul Rosenbloom
University of Southern California
Information Sciences Institute and
Computer Science Department
4676 Admiralty Way
Marina del Rey, CA 90292, USA

…Knowing how to create compelling experiences; do low-cost, high-performance computing; support large-scale network simulations; build graphics-modeling software is (or will be) [the entertainment community] stock and trade. In these areas not only will it be futile for the Army to try to compete, but a waste of energy and resources. Bran Ferren [1]

Introduction

Bran Ferren makes a compelling argument that the Entertainment Industry is driving the technology advances needed for military virtual reality systems. Moreover, the military virtual environment community may actually be falling behind its civilian counterparts by ignoring the rapid changes going on in entertainment computing. These advances include low cost computer graphics, agent technology, and the use of 3D audio.

In this paper we will explore some of the reasons how and why the Entertainment Industry is advancing the state of virtual reality (VR). We will also look at the current problems of military simulation, particularly its lack of story and emotion. Finally, this paper examines how the US Army is trying to address these issues with the establishment of the Institute for Creative Technology (ICT).

The Entertainment Industry

The Entertainment Industry has in many ways grown far beyond its military counterpart in influence, capabilities and investments. For example, Microsoft alone expects to increase R&D spending next year by 23 percent, to \$3.8 billion — compared to the US Army's \$ 1.2 billion science and technology budget. The Interactive Digital Software Association estimates that in 1998, interactive entertainment businesses invested approximately \$2 billion in new technology R&D, with an increase of more than 20 percent. [2] This far outweighs current US Army research and development for training and simulation technology.

Moreover, the advances in the industry cannot be ignored. Witness the rapid pace of development of the graphics systems for game consoles and personal computers — almost double performance every nine months [3]. Compare this with the relatively slow gains in "high-end" graphics platforms being used for the military.

According to Richard Weinberg at the University of Southern California's School of Cinema-Television, Sony's upcoming PlayStation 2 is an example of a consumer-grade advanced technology gaming platform that could revolutionize both the world of home gaming as well as interactive training for the Army. The PS2 is expected to have 34 times the power of the current leading game system, the Sony PlayStation, and more than twice the graphics performance of SGI's (formerly Silicon Graphics) high-end visualization system, the Infinite Reality 2. Here is what *Game Informer Magazine* (May 1999) says about the upcoming Playstation 2: "PlayStation 2 could be a glimpse at Hollywood of the 21st Century. Developers with this kind of power in their hands could theoretically create real-world environments, with living breathing characters all affected by real-world physical attributes such as gravity, friction and mass. Plus, PS2 can accurately simulate different materials such as water, wood, metal, and gas — real worlds that look like real worlds. Full motion video that's not full motion video, but real-time game play with speaking characters, fluid motions, and facial expressions."

Playstation 2Graphics Synthesizer – Features and General Specifications:

- GS Core: Parallel Rendering Processor with embedded DRAM
- Clock Frequency: 150 MHz
- No. of Pixel Engines: 16 (in Parallel)
- Embedded DRAM: 4 MB of multi-port DRAM (Synced at 150MHz)
- Total Memory Bandwidth: 48 gigabytes per second
- Combined Internal Data Bus Bandwidth: 2,560 bit
- Read: 1,024 bit
- Write: 1,024 bit
- Texture: 512 bit
- Display Color Depth: 32 bit (RGBA: 8 bits each)
- Z Buffering: 32 bit
- Rendering Functions: Texture Mapping, Bump Mapping, Fogging, Alpha Blending, Bi- and Tri-Linear Filtering, MIPMAP, Anti-aliasing, Multi-pass Rendering

[1] For contact with authors: michael_macedonia@stricom.army.mil, and Rosenbloom@isi.edu

Paper presented at the RTO HFM Workshop on "What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?", held in The Hague, The Netherlands, 13-15 April 2000, and published in RTO MP-058.

Rendering Performance:

- Pixel Fill Rate: 2.4 giga pixel per second (with Z buffer and Alphablend enabled), 1.2
- giga pixel per second (with Z buffer, Alpha and Texture)
- Particle Drawing Rate: 150 million/sec
- Polygon Drawing Rate: 75 million/sec (small polygon), 50 million/sec (48 pixel quad with Z and A), 30 million/sec (50 pixel triangle with Z and A), 25 million/sec (48 pixel quad with Z, A and T)
- Sprite Drawing Rate: 18.75 million (8 × 8 pixels)

Digital Output:

- NTSC/PAL
- Digital TV (DTV)
- VESA (maximum 1280 × 1024 pixels)

Other technical trends that will likely shape the military training world will be digital cinema, the convergence of television with the World Wide Web, and the continued rapid growth of multiplayer Internet 3D games such as Sony's *Everquest*.

Weinberg also notes that, from a content perspective, the computer game industry has considerable expertise in games relevant to aspects of military training, with significant interest in war games, simulations, and military-like shooter games. For example, TalonSoft's *The Operational Art of War II* is expected to cover the Vietnam War, Arab-Israeli wars, the Iran-Iraq conflict, and Operation Desert Storm at the operational command level, as well as several hypothetical conflict scenarios ranging from India/Pakistan to a new Korean conflict. *Extreme Tactics*, *Warbreeds*, and *WarZone 2100* are but a few examples of the war/strategy/shooter-style games available. According to the May 15, 1999 issue of *Games Business*, PC games by genre, ranked by unit share from April 1998-March 1999 were comprised of Strategy 21.8%, Simulation 13.4%, Adventure/role playing 12.1% and Action 11.4%. [2]

Even traditional flight simulation companies are taking advantage of the emergence of commercial game software for training. For example, Flight Safety International re-markets a version of Microsoft Flight Simulator and the Navy is experimenting with the game for new pilot training.

What's Wrong with Military VR?

Until recently, the military has led the way in developing advanced virtual environments. We know the importance of experiential learning through the development and use of the National Training Center, Conduct of Fire Trainers, Simnet, and flight simulators. The vision of the military VR community has been to develop realistic virtual environments to support training, mission rehearsal, concept exploration and engineering design.

However, military simulations currently fall short of enabling this vision of realism for a multitude of reasons. First, the necessary technology does not yet exist, and must be created. Our ability to immerse participants is quite limited. For example, with respect to physical immersion, it is currently possible to provide good auditory, moderate visual, and primitive tactile/haptics while essentially no olfactory or gustatory immersion is possible. The ability to track full body motion, gesture and expression is still nascent while virtual mobility is limited to primitive two-dimensional approaches.

What technologies do exist for physical immersion tend to be neither portable nor wireless. They also have interoperability problems, fail to scale well to large numbers of entities and have latency problems when it comes to closely coupled interactions over long distances. Defining (modeling), organizing and distributing multimedia content also can be a problem.

Second, the stories and characters used in military simulations are skeletal and rudimentary. A typical story consists of a background briefing plus an event list. A typical character is defined in terms of a role and a set of scripted behaviors. Some degree of intellectual immersion, to the extent of triggering some of the same key decision making tasks that would occur in the real world, is possible with such minimal story and character definitions. However, rich story and engaging characters can more fully engross the participant and provide a more appropriate context for intellectual activity. (Note that for peacekeeping training the US Army often hires actors for live exercises at its Combat Training Centers.)

Lack of rich story and character also impairs emotional immersion, as abstractions do not generally induce intense emotions. Because emotions are powerful motivators, and can lead to significant shifts in both how the world is interpreted and how decisions are made — in the extreme, it can be a matter of decision making in a life-or-death situation — this lack of emotional immersion is a major gap in making realistic simulations. Emotional immersion is a particular strength of the entertainment industry.

Third, the full set of necessary people to solve these problems has been incomplete. Technical personnel working with domain experts currently build military simulations. This collaboration is critical, but creative personnel — such as writers and cinematographers — need to be added to the mix. The further advantages of such a combination are that technical advances can open up new creative realms, creative needs can drive new research, and creative techniques can mask limitations in technology.

[2] Weinberg was a key member in the development of the ICT proposal.

Recognition

Early in 1999, US Army leaders recognized a need for a major transformation of our forces and the limitations of our current simulation technologies. Furthermore, this transformation required the ability to develop new training and simulation systems for future conflicts that leveraged the capabilities of both the entertainment industry and academia.

The US Army and Department of Defense selected the University of Southern California (USC) as a strategic partner in the development of the Institute for Creative Technologies (ICT) because of its unique confluence of scientific capabilities and Entertainment Industry relationships necessary for leadership in simulation.

The prime objective, as reaffirmed by Dr. Michael Andrews, Deputy Assistant Secretary of the Army for Research and Technology, was to build a special partnership with the entertainment industry and academia. Furthermore it was to advance the state of the Army's technology and transition it quickly to programs such as the Future Combat System.

A University Affiliated Research Center (UARC) is a strategic relationship, requiring both breadth and depth in capabilities matched with industry partnership to achieve major advancements in science and technology.

This model of research is not new. For example The National Automotive Center (NAC) serves as the Army's focal point for the development of dual-needs/dual-use automotive technologies and their application to military ground vehicles. The NAC identifies the common needs of the Defense Department, automotive industry and academia for the purpose of collaborative research and development.

Part of USC’s uniqueness arises from its location in Los Angeles, at the hub of both the entertainment and aerospace industries; part arises from its standing as a leading private research university; and part arises from the capabilities and stature of its component units, and the working relationships they have developed with industry.

USC’s top-ranked School of Cinema-Television grew up with the entertainment industry and continues to maintain uniquely close ties with it. USC’s School of Engineering is ranked 12th in the nation. Its Information Sciences Institute is home to leading academic research groups in networking and artificial intelligence. USC's top-ten (and in some rankings, top-five) ranked Annenberg School for Communication leverages off of the Los Angeles area's varied strengths in new technology, telecommunications, film, television, radio, newspapers and magazines, and policy and research organizations.

The Institute for Creative Technology

USC established ICT under the auspices of the US Army Simulation, Training, and Instrumentation Command (STRICOM) to focus on developing the art and technology for synthetic experiences that are so compelling participants will react as if they are real. That is, ICT will bring *verisimilitude* — the quality or state of appearing to be true — to synthetic experiences. This will produce a revolution in how the military trains and how it rehearses for upcoming missions; just to name two quite specific, but highly critical, military needs. However, more generally, it will provide a quantum leap in helping the Army prepare for the world, soldier, organization, weaponry, and mission of the future. Beyond the military, ICT will also advance a compelling new medium for (at least) entertainment, education, arts, and travel.

From the start, ICT leverages heavily off of this dual-use nature by actively engaging the Entertainment Industry (comprising film, TV, interactive gaming, etc.) and possibly other industries later. ICT will serve as a means for the military to learn about, and benefit from, the technologies that are being developed in the Entertainment Industry, and for transferring technologies from the Entertainment Industry into the military. ICT will also work with creative talent from the Entertainment Industry in order to adapt their concepts of story and character to increasing the degree of immersion experienced by participants in synthetic experiences, and to improving the utility of the outcomes of these experiences.

ICT will pursue a combination of basic and applied research (plus some educational activities). Basic research will cover six thrusts crucial to the kind of verisimilitude that is the institute’s mission [4]:

1. Immersion — Providing compellingly realistic experiences
2. Networking and Databases — Organizing, storing and distributing content
3. Story — Providing compelling interactive narratives that propel experiences
4. Characters — Replacing human participants with automated ones
5. Setup — Authoring and initializing environments, models and experiences
6. Direction — Monitoring, directing, and understanding experiences

Applied research will be organized around a small number of long-term themes; for example, simulating futuristic style forces. Within each theme, a set of key projects will be identified, along with an integration architecture that will eventually bring them all together in a single system covering the theme. Projects will be pursued via sequences of prototypes of increasing functionality and level of integration. The Army and the Entertainment Industry will be actively involved at each

step in helping to ensure that what is done meets their needs.

Key elements associated with USC's array of relevant existing capabilities include:

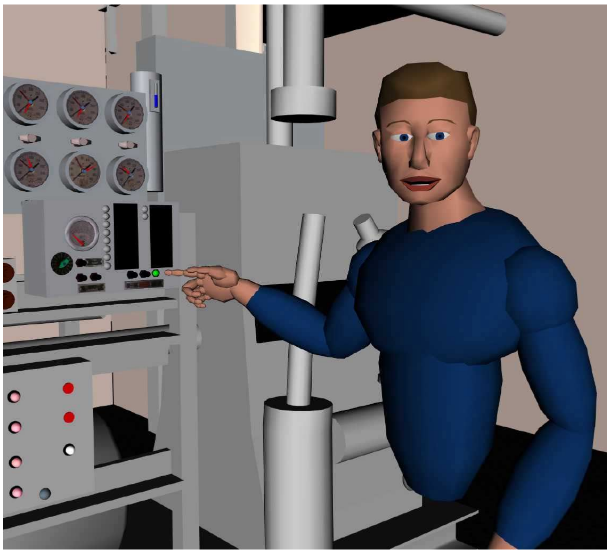
- The Entertainment Technology Center, (ETC) which is a research and development project of the School of Cinema-Television. ETC's mission is to discover, research, develop and accelerate entertainment technology. Steven Spielberg and George Lucas sit on the ETC board.
- The Annenberg Center for Communication that advances communication and information technologies through interdisciplinary research and outreach.
- The Integrated Media Systems Center, (IMSC), a National Science Foundation (NSF) established center providing multi-media technologies. USC successfully outbid 117 other university competitors in response to the 1996 NSF national competition for an integrated media center.
- The Information Sciences Institute, which combines world class research and development across a broad range of computer science and engineering with a strong relationship with the Department of Defense.

ICT Vision

The vision for the ICT is to develop the art and technology for synthetic experiences that are so compelling participants will react as if they are real. Participants will be fully immersed physically, intellectually, and emotionally. They will be capable of full three-dimensional mobility. Their behavior will be propelled through engrossing stories stocked with engaging characters that may be either automated or manned — the high quality of the automated characters along with the provision of plug compatibility will make it impossible to distinguish. They will interact with the experiences as if they are real. In short, the ICT will provide a new meaning for "high fidelity": *verisimilitude*.

Imagine the soldier of the not so distant future. It is Sunday and he is at home in Los Angeles. He and his best friend in Hong Kong are relaxing by immersing themselves in the nostalgic world of the 1990s. They are founding an Internet startup company during the heyday of the speculative bubble, learning to deal with venture capitalists, trying to fend off large predatory rivals, and ultimately trying to steer their new company towards a successful Initial Public Offering. However, just when the story is getting really engrossing, a high priority videomessage arrives from his commanding officer with the news that he will be shipping out within a few days, along with the five thousand or so other members of his Strike Force.

The mission will be to help keep the peace in the latest global hot spot, but there are not yet any details concerning his unit's specific mission or the volatile situation that currently exists on the ground there. He also knows nothing about the country's history, culture or language. Fortunately he has a long flight ahead of him, and the Army is ready for him.



STEVE is an intelligent tutor developed by USC/ISI for the Office of Naval Research. [6]

He begins the flight with a brief on-line course covering the history and culture of the region. A virtual tutor helps him make the best possible use of the very limited time he has available. (See figure). He then dons his personal immersion system and walks into a simulated market in the capital city, where a helpful (computer generated) shopkeeper introduces him to the basic aspects of the language along with the range of interpersonal interaction styles — both positive and negative — common to the culture.

Next, he is briefed by his commanding officer on his unit's mission — to keep innocent civilians from being hurt in factional violence while preventing, as much as possible, new flare ups among the factions. By sharing an immersive space with his commander and the rest of his unit — even though in reality they are physically dispersed across several transport aircraft — he is able to join them for a quick tour of their area of responsibility, followed by a session in which they are able to familiarize themselves with the uniforms and weapons used by the various factions. He can pick up the uniforms and examine them as well as see them on various models. He can try out the weapons himself, as well as pull up specs and performance numbers on them. At all times he can discuss what he sees and does with his commander and the other members of his unit.

During his final few hours he is immersed in a sample mission. The sights, sounds and smells of the city immediately bombard him. There are people everywhere going about their lives as best they can. He's a bit scared and hesitant at first, but fortunately the rest of his unit is there in the street with him. There's a second unit

nearby, however he is unaware that they — along with all of the citizens with whom he is interacting — are computer-generated characters.

He is in a large central plaza in the city. A bazaar is located in one part of the plaza and throngs of people are milling about bartering for various goods. The plaza is ringed by several government buildings and at the far end there is a large church. The scene is a rich and confusing tapestry of life — our soldier struggles to remember the identifying features of the various factions as he attempts to make sense of the scene. Suddenly, near the church, a large disruption occurs and reports ring out, echoing off the buildings. What is going on? Is one of the rebel factions trying to attack the government? Rifles at the ready, he and other members of his squad rush toward the disturbance, where they confront — a wedding party leaving the church and a group of celebrants setting off large firecrackers.

Switching the safety back on, he shoulders his rifle and breathes a sigh of relief while a computer generated tutor emphasizes the need to assess the situation before taking action and points out that in this culture celebrations are often accompanied by fireworks which can be mistaken for gunfire. This kind of immediate feedback is enabled through the use of computer agents as tutors. Because it is provided in context, it can be much more effective than an after action review, where there may be a substantial delay between the exercise and the review.[3]

This scenario was orchestrated by the Director, another computer agent that directs the behavior of the other agents in the simulation and the environment. By exercising control of these elements, the Director ensures that the exercise follows the intended story line so that the intended training goals can be achieved. In this case, this scenario was intended to create a situation in which the soldier would be confronted with an ambiguous but potentially threatening situation where it would be necessary to decide whether or not to act — and where the wrong decision would have disastrous consequences.

Although the soldier in the exercise is free to make choices, the Director manipulates the simulation so that eventually he is forced to confront the intended dilemma, thereby achieving the pedagogical goals for the simulation. For example, if the soldier and his squad had not noticed the initial disturbance, the wedding celebration would have become louder and more boisterous, until it could not be ignored. Furthermore, the squad's failure to recognize the disturbance in its early stages would be an issue that the tutor would cover during its *in situ* review of the exercise.

This is just one of many possible examples of the kind of experience that ICT will make possible and, in fact, commonplace. Verisimilitude of this sort will require combining the art of (interactive) storytelling with the art and technology of transforming these stories into compelling interactive experiences. It inherently involves collaboration between the kinds of creative and technical experts found in the entertainment industry and the kinds of researchers and system builders found in the academic, industrial and military R&D communities. Fortunately, all of these necessary partners are either already present at USC or linked closely with it.

We expect that by creating a true synthesis of art and technology[4] and of the capabilities of the entertainment industry and the R&D community — all in service of verisimilitude — military training and mission rehearsal will be revolutionized by making it more effective in terms of cost, time, the types of experiences that can be trained or rehearsed, and the quality of the result. It will also provide a new medium for entertainment, enabling both individuals and groups to be fully immersed and engaged in compelling experiences from their homes, or wherever they happen to be located.

Beyond entertainment, verisimilitude will also provide new media for (at least) both immersive distance learning and the arts (particularly the performing arts). It could also even support a new mode of virtual travel; providing immersive presence in a remote location, and augmenting the local populace (with whom direct interaction may not be possible) with synthetic characters with whom interaction is possible.

Conclusion

The computer and Internet revolutions have substantially changed the direction of entertainment from delivery in a mass medium such as television to a mass customized experience via the Web and PC. However, the art of entertainment still requires stories, characters and direction to make the experience meaningful and enjoyable.

The US Army faces the same challenge of adapting to the changes brought about through the mass marketing of supercomputing (e.g. Playstation 2), low-cost graphics, and the higher expectations of technically savvy soldiers.

Moreover, a more fundamental need is to represent new kinds of problems such as urban conflict, operations other than war, and information operations that cannot be simulated well in military virtual environments today. As the vignette presented above demonstrated, there is an urgent requirement to represent the human dimensions of war and conflict to provide training for the truly difficult decision-making problems our soldiers

[3] This vignette was partly developed by William Swartout, the Technical Director for ICT.

[4] Providing what Richard Lindheim, the Executive Director of ICT, has referred to as *Show Technology* as a complement to the more common combination of art and business as *Show Business.*

must face. NATO's experience in Kosovo is now a common one for countries such as the United Kingdom (e.g. Northern Ireland).

The establishment of the Institute for Creative Technologies is just one of many steps needed to providing the essence of verisimilitude into training and virtual reality systems. The US Army will explore all avenues of entertainment technology to keep pace with the challenges presented to us, whether in application to distributed learning or embedded training systems. Ultimately, we want to prepare our soldiers for the future by experiencing it.

References

[1] Bran Ferren, Some Observations on the Future of Army Simulation, *RDA Magazine*, May 1999, US Army, Washington, DC, pp. 31-37.
[2] Michael Macedonia, Why Digital Entertainment Drives the Need for Speed, *IEEE Computer*, February 2000, pp. 124-127.
[3] Michael Macedonia, Sony versus Wintel: Mortal Combat, *IEEE Computer*, July 1999, p. 112.
[4] Michael Zyda, ed., *Modeling and Simulation: Linking Entertainment and Defense*, National Research Council, National Academy of Sciences Press, Washington, DC 1997.
[5] Nat Durlach and Anne Mavor, eds., Committee on Virtual Reality Research and Development, *Virtual Reality: Scientific and Technological Challenges*, National Research Council, National Academy of Sciences Press, Washington, DC 1994.
[6] J. Rickel and W.L. Johnson. Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor Control. *Applied Artificial Intelligence* pp. 13:343-382, 1999.

Biographies

Paul Rosenbloom is currently Professor in the Computer Science Department at the University of Southern California (USC), New Directions at the Information Sciences Institute (USC/ISI), and Deputy Director of the Intelligent Systems Division at USC/ISI. Prior to coming to USC in 1987, was an Assistant Professor of Computer Science and Psychology at Stanford University from 1984 to1987, and a Research Computer Scientist at Carnegie Mellon University from 1983 to 1984. Received a B.S. degree in Mathematical Sciences from Stanford University in 1976 and M.S. and Ph.D. degrees in Computer Science from Carnegie Mellon University in 1978 and 1983, respectively.

Michael Macedonia is chief scientist and technical director of the US Army Simulation, Training, and Instrumentation Command, Orlando, Fla. A graduate of West Point, Macedonia served as an infantry officer in a number of United States and overseas assignments including Germany and the Middle East. He is a veteran of Desert Storm. Following his military service, Macedonia became the Vice-president of the non-profit Fraunhofer Center for Research in Computer Graphics, Inc. (CRCG) in Providence, Rhode Island. Macedonia then joined the Institute for Defense Analyses in Alexandria, Virginia. His M.S. is in Telecommunications from the University of Pittsburgh and he has a Ph.D. degree in Computer Science from the Naval Postgraduate School.

Information on ICT

For more information on ICT, please contact the Program Manager, Dr. James Blake, at STRICOM. His email is james_blake@stricom.army.mil or +1 407 384-3502.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-MP-058 AC/323(HFM-058)TP/30	ISBN 92-837-1057-6	UNCLASSIFIED/ UNLIMITED

5. Originator	Research and Technology Organization North Atlantic Treaty Organization BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France
6. Title	What Is Essential for Virtual Reality Systems to Meet Military Human Performance Goals?
7. Presented at/sponsored by	the Workshop of the RTO Human Factors and Medicine Panel (HFM) held in The Hague, The Netherlands, 13-15 April 2000.

8. Author(s)/Editor(s)	9. Date
Multiple	March 2001
10. Author's/Editor's Address	**11. Pages**
Multiple	172

12. Distribution Statement	There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Military training	Performance
Human factors engineering	Humans
User needs	Man machine systems
Interfaces	Perception
Research projects	Virtual reality
Motion sickness	Virtual environments
Similators	

14. Abstract

This workshop aimed to identify the functional requirements of potential military applications of Virtual Reality (VR) technology, to report the state-of-the-art and projected capabilities of VR technologies, and to propose future research requirements and directions for military applications.

During the workshop discussions, forty participants from military organisations, academia and industry put forward their opinions on the significant bottlenecks and opportunities in the development of military VR applications. Presentations discussed visual, haptic, auditory and motion feedback, navigation interfaces, and scenario generation, modelling software and rendering hardware.

VR research transition opportunities include the domains of training, planning & mission rehearsal, simulation supported operation, remotely operated systems and product design.

Critical bottlenecks are a lack of natural interfaces, a lack of technology standardisation and a lack of behavioural models and team interaction tools in VR.

In general, better co-ordination between military organisations, industry and academia is necessary in order to identify gaps in current knowledge and to co-ordinate research. Suggestions for closing gaps are included.

This page has been deliberately left blank

Page intentionnellement blanche

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25 • 7 RUE ANCELLE
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS
RTO NON CLASSIFIEES

L'Organisation pour la recherche et la technologie de l'OTAN (RTO), détient un stock limité de certaines de ses publications récentes, ainsi que de celles de l'ancien AGARD (Groupe consultatif pour la recherche et les réalisations aérospatiales de l'OTAN). Celles-ci pourront éventuellement être obtenues sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous par lettre ou par télécopie à l'adresse indiquée ci-dessus. Veuillez ne pas téléphoner.

Des exemplaires supplémentaires peuvent parfois être obtenus auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus sur la liste d'envoi de l'un de ces centres.

Les publications de la RTO et de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr, (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

BELGIQUE
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
Directeur - Recherche et développement - Communications et gestion de l'information - DRDCGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)
Hellenic Ministry of National Defence
Defence Industry Research & Technology General Directorate
Technological R&D Directorate
D.Soutsou 40, GR-11521, Athens

HONGRIE
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Centro di Documentazione Tecnico-Scientifica della Difesa
Via XX Settembre 123a
00187 Roma

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

PAYS-BAS
NDRCC
DGM/DWOO
P.O. Box 20701
2500 ES Den Haag

POLOGNE
Chief of International Cooperation Division
Research & Development Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
Distribuční a informační středisko R&T
VTÚL a PVO Praha
Mladoboleslavská ul.
197 06 Praha 9-Kbely AFB

ROYAUME-UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
Etats-Unis

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume-Uni

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery
Montreal Road, Building M-55
Ottawa K1A 0S2, Canada

Les demandes de documents RTO ou AGARD doivent comporter la dénomination "RTO" ou "AGARD" selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
Etats-Unis

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
Etats-Unis
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par St-Joseph Ottawa/Hull
(Membre de la Corporation St-Joseph)
45, boul. Sacré-Cœur, Hull (Québec), Canada J8X 1C6

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Telefax 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

NATO's Research and Technology Organization (RTO) holds limited quantities of some of its recent publications and those of the former AGARD (Advisory Group for Aerospace Research & Development of NATO), and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. **Please do not telephone**.

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO publications, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your organisation) in their distribution.

RTO and AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
Director Research & Development Communications & Information Management - DRDCIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
Distribuční a informační středisko R&T
VTÚL a PVO Praha
Mladoboleslavská ul.
197 06 Praha 9-Kbely AFB

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (ISP)
29 Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr, (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

GREECE (Point of Contact)
Hellenic Ministry of National Defence
Defence Industry Research & Technology General Directorate
Technological R&D Directorate
D.Soutsou 40, GR-11521, Athens

HUNGARY
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Centro di Documentazione Tecnico-Scientifica della Difesa
Via XX Settembre 123a
00187 Roma

LUXEMBOURG
See Belgium

NETHERLANDS
NDRCC
DGM/DWOO
P.O. Box 20701
2500 ES Den Haag

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

POLAND
Chief of International Cooperation Division
Research & Development Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
United States

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Document Delivery
Montreal Road, Building M-55
Ottawa K1A 0S2, Canada

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
United States

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by St. Joseph Ottawa/Hull
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada J8X 1C6

ISBN 92-837-1057-6